Rend. Sem. Mat. Univ. Pol. Torino - Vol. 64, 2 (2006)

D. A. Bini

NUMERICAL SOLUTION OF LARGE MARKOV CHAINS*

Abstract. Markov chains model a wide variety of problems of the real world where a given system evolves in time and can assume different states according to the occurrence of some random events. Solving large Markov chains is often a challenging problem for the large complexity encountered by the available algorithms. This is a typical case where the technological advances of the computer hardware are not enough for providing a solution in almost real time unless very specific and highly specialized algorithms are designed. The advances of the mathematical technology is the key tool for solving such problems.

In this paper we rely on two examples, the Google problem and the shortest queue problem, used as paradigms in order to show the main techniques and the mathematical tools for the solution of large Markov chains. The emphasis is set on the role of the mathematical structures which characterize the problem and on their analysis and exploitation. We present an algebraic version of the Ramaswami formula for solving infinite Markov chains associated with queueing problems. Then we revisit it from a functional point of view in terms of canonical Wiener-Hopf factorization of a matrix valued function of complex variable. We show that computing this factorization is reduced to solving a matrix equation or to inverting a matrix Laurent power series. Classical and more recent advanced algorithms for the latter two computations are presented.

1. Introduction

Markov chains model a wide variety of problems of the real world where a given system evolves in time and can assume different states according to the occurrence of some random events. There are many examples of Markov chains where for the huge dimensions encountered in practice the solution of the system leads to challenging computational problems in linear algebra. Besides the classical examples of the *random walk* problems, a great interest is addressed to a large part of *queueing models* which are encountered in telecommunications and are related to the web technology. In particular, a special attention has recently received the *page rank problem* carried to the attention of researchers by the Google$^{\text{TM}}$ search engine where the dominant eigenvector of a matrix of size about 8.5×10^9 must be computed in almost real time.

In this paper we address the attention on computational problems related to the solution of large Markov chains by pointing out the fundamental importance of the theoretical and algorithmic tools at the basis of their solution. These tools make solvable, in almost real time, complex problems which would be untreatable with a classical numerical approach even with the fastest computers available nowadays.

The paper is organized as follows. In Section 1.1 we informally recall the elementary concepts at the basis of Markov chains; in Section 1.2 we present two important examples: the page rank (Google) problem and the shortest queue problem modeled by huge and infinite Markov chains, respectively. In Section 2 some computational issues are introduced: here the attention is addressed to the role of structure analysis

*"LEZIONE LAGRANGIANA" given on November 23rd, 2005.

for the design of highly efficient algorithms. Section 3 deals with computational tools used for solving infinite Markov chains, in particular the Ramaswami formula and its relationships with certain factorizations of matrix valued functions in the Wiener algebra. It is shown that solving an infinite Markov chain of the M/G/1 type is reduced to computing a Wiener-Hopf factorization of a suitable matrix valued function defined on the complex unit circle. The algorithmic aspects of this kind of factorizations are considered in Section 4. It is pointed out that a Wiener-Hopf factorization is readily available if the solution with minimum spectral radius of a suitable matrix equation is computed, or if the matrix inverse of a matrix Laurent power series defined on the unit circle is computed. Sections 5, 6 and 7 deal with these two latter computations. In particular, in Section 7 we describe the algorithmic advances obtained in solving matrix equations by relying on the cyclic reduction algorithm which is revisited from a functional point of view. Conclusions and open problems are the subject of Section 8.

A particular emphasis is set on the analysis of the mathematical structures which characterize the problem and on the importance of looking at the problem itself from different point of view. In fact, the latter approach enables one to have a more complete understanding of all the theoretical and computational features encountered in the analysis.

1.1. Markov chains

Informally speaking a Markov chain is a sequence $\{X_k\}_{k\in T}$ of random variables indexed on a discrete set T representing time, where each variable X_k can take values in a set E of *states* this value represents the status of the system at time k [16]. Usually it is assumed that $T = \mathbb{N}$. The fundamental property of a Markov chain is that the state occupied by the system at time k only depends on the state occupied by the system at time $k-1$. That is, the value of X_k only depends on the value taken by X_{k-1} and is *independent of* the past history of the system.

This property can be formalized by the following equation

$$\mathtt{P}[X_k = i_k | X_0 = i_0, \ldots, X_{k-1} = i_{k-1}] = \mathtt{P}[X_k = i_k | X_{k-1} = i_{k-1}]$$

for any $i_0, \ldots, i_k \in E$, where $\mathtt{P}[X = a | Y = b]$ denotes the conditional probability that the random variable X takes the value a given that the random variable Y takes the value b.

The evolution of a Markov chain is determined by the probability $p_{i,j}^{(k)}$ that at each time k the system moves from state i to state j. If this array of probabilities is independent of the time k, say, $p_{i,j}^{(k)} = p_{i,j}$, then the chain is called *homogeneous*. Thus, a homogeneous Markov chain is characterized by the *transition probability matrix* $P = (p_{i,j})_{i,j\in E}$. Observe that P is *row stochastic*, that is, $p_{i,j} \geq 0$ and $\sum_{j\in E} p_{i,j} = 1$ for any $i \in E$. In matrix form one has $P\mathbf{e} = \mathbf{e}$ where $\mathbf{e}$ is the vector indexed on E with components equal to 1.

The status of the system at time k is described by a *probability state vector* $\mathbf{x}^{(k)} = (x_i^{(k)})_{i\in E}$ where $x_i^{(k)}$ denotes the probability $\mathtt{P}[X_k = i | X_0]$ that the system is in the state $i \in E$ at time k.

Indeed, by the lows of probability one has

$$(\mathbf{x}^{(k+1)})^T = (\mathbf{x}^{(k)})^T P \tag{1}$$

and $\sum_{i \in E} x_i^{(k)} = 1$.

Equation (1) describes the evolution of the Markov chain once the initial probability state vector $\mathbf{x}^{(0)}$ has been assigned.

Observe that, if the system has an asymptotic behavior, i.e., if there exists the limit $\lim_k \mathbf{x}^{(k)} = \pi$ then from (1) one has

$$\pi^T = \pi^T P$$

that is, π^T is a nonnegative left eigenvector of P corresponding to the eigenvalue 1 normalized so that $\pi^T \mathbf{e} = 1$.

The vector π if it exists is called *invariant probability vector* or *steady state vector*.

One of the the main computational issues in the solution of Markov chains is the computation of π.

1.2. Some examples

A classical example of Markov chain is the *random walk* problem. Consider a system made up by a single particle which may occupy positions in the segment $[1, n]$ given by the integer coordinates $i = 1, 2, \dots, n$. The particle can move to the right or to the left with probability p and $1-p$, respectively, of a unit step unless it occupies one edge of the segment. In this case the particle moves with probability 1 of one step toward the center of the segment. The system, known as random walk, is described by a Markov chain where the set of the states is $E = \{1, 2, \dots, n\}$ and

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1-p & 0 & p & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1-p & 0 & p \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}.$$

If the set $\{1, 2, \dots, n\}$ is replaced by $\mathbb{N}$, the particle moves along a half-line, and P turns into the semi-infinite matrix

$$P = \begin{bmatrix} 0 & 1 & 0 & \cdots \\ 1-p & 0 & p & \ddots \\ 0 & 1-p & 0 & p \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

If $E = \mathbb{Z}$, the particle moves along a line, and P is a bi-infinite matrix.

The *page rank problem* [1, 14, 29, 36, 40, 44] consists in sorting a set of web pages according to their relevance. This problem is crucial for search engines like *Google*^{TM} widely used by people surfing on the web. Google treats this problem by means of a Markov chain simulating a virtual user who randomly surfs on the web moving from one page to another one according to the following rules:

– Number the pages available in the web from 1 to n, that is, set $E = \{1, 2, \ldots, n\}$; divide the time into intervals of length 1.

– At the end of each time interval the virtual surfer, while visiting the page i decides to change page and has two different choices

a) with probability θ he/she randomly chooses a link available in the current page and moves there;

b) with probability $1 - \theta$ he/she randomly chooses a new available page in the web, i.e., a page from the set E

Assume that in his/her random choice the virtual surfer chooses with uniform probability among the links of the current page (case a) and among the available pages (case b). Then the probability transition matrix P is given by

$$P = \theta DH + \frac{1}{n}(1-\theta)\mathbf{e}\mathbf{e}^T$$

where $H = (h_{i,j})$, called the *connectivity matrix* is such that $h_{i,j} = 1$ if in the page i there is a link to the page j, $D = \text{Diag}(d_1^{-1}, \ldots, d_n^{-1})$ with $d_i = \sum_{j \in E} h_{i,j}$, where we assume that $d_i \neq 0$ for any i. Indeed, the matrix θDH describes the case a) while $\frac{1}{n}(1-\theta)\mathbf{e}\mathbf{e}^T$ describes the case b).

For $0 \leq \theta < 1$ the invariant probability vector $\pi = (\pi_i)_{i=1:n}$ of this Markov chain exists and is unique in view of the Perron-Frobenius theorem since P is irreducible [47]. Moreover, π_i is the limiting probability that the surfer visits the ith page. The higher is the probability of a page to be visited the higher is the importance of the page. Computing the vector π provides a reliable way of ranking the web pages. Google makes this computation once in a month. At any query, the list of selected pages is sorted according to the order provided by π.

It is interesting to point out that at the date of December 2005 there are roughly $n = 8.5 \cdot 10^9$ pages available in the web. The matrix P, also called Google matrix, has size n and $n^2 = 7.2 \cdot 10^{19}$ elements. This matrix is probably the largest finite matrix ever encountered in applications.

There are important applications where the Markov chain is described by an *infinite matrix*. This is the case of queueing problems. Consider for instance the following simple model problem describing the queues which any driver encounters at the highway exits (see figure 1).

In the *shortest queue problem* there are m servers which provide some service to customers upon their arrival. Time is discretized into intervals of length one, during each interval α customers arrive and each of them joins the shortest line; at the end of

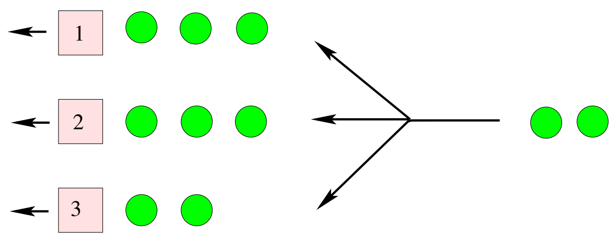

Figure 1: The shortest queue problem: m servers provide a service to m customers who leave the lines; new customers arrive and join the shortest queue.

each interval each server provides the service and m customers, if there are any, leave the system. We assume that the distribution of the random variable α is known, i.e., the probabilities $q_i = \mathtt{P}[\alpha = i]$, for $i = 0, 1, \ldots$ are known. Denoting by X_n the overall number of customers waiting to be served at time n it holds

$$X_{n+1} = \max\{0, X_n - n + \alpha\}$$

and the probability transition matrix is readily available:

$$p_{i,j} = \begin{cases} q_0 + q_1 + \cdots + q_{m-i} & \text{if } j = 0, \;\; 0 \le i \le m-1 \\ q_{j-i+m} & \text{if } j - i + m \ge 0 \\ 0 & \text{if } j - i + m < 0 \end{cases}$$

Since we assume no upper bound to the length of the queues, the set E of states coincides with $\mathbb{N}$ and the matrix P is semi-infinite

$$P = \begin{bmatrix} b_m & q_{m+1} & q_{m+2} & q_{m+3} & \cdots & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\ b_1 & q_2 & q_3 & q_4 & \ddots & \ddots \\ q_0 & q_1 & q_2 & q_3 & \ddots & \ddots \\ & q_0 & q_1 & q_2 & q_3 & \ddots \\ O & & \ddots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{2}$$

where $b_i = q_0 + q_1 + \cdots + q_i$, $1 \le i \le m$.

For this problem, as well as in any other queueing model, knowing the components of π provides information about the expected length of the queues. Since $\sum_{i \in E} \pi_i = 1$ and $\pi_i \ge 0$, the values π_i decay to zero and the computation of a finite number of components of π, say k, provides global information about the remainder since $\sum_{i>k} \pi_i = 1 - \sum_{i=0}^{k} \pi$.

Observe that, by partitioning the matrix P into $m \times m$ blocks, one can represent

P in block Hessenberg form as

$$P = \begin{bmatrix} \widehat{A}_0 & A_1 & A_2 & \dots \\ A_{-1} & A_0 & A_1 & \ddots \\ & A_{-1} & A_0 & \ddots \\ O & & \ddots & \ddots \end{bmatrix} \tag{3}$$

where $A_k = (q_{m+j-i})$ and $\widehat{A}_0$ is obtained by replacing the first column of A_0 with $(b_m, \dots, b_1)^T$.

Observe also that except for the block in position (1, 1), the blocks along each diagonal coincide. Matrices having constant blocks along their block diagonals are called *block Toeplitz*. The fact that $\widehat{A}_0 \neq A_0$, which expresses the "boundary condition" of the Markov chain, is needed to maintain the stochasticity of the matrix P.

Infinite block Toeplitz matrices, except for the boundary condition, in block upper Hessenberg form describe a much more general class of Markov chains called M/G/1 [37, 38]. Infinite block Toeplitz matrices, except for the boundary condition, in lower Hessenberg form, describe the class of G/M/1 Markov chains. Infinite block tridiagonal block Toeplitz matrices, except for the boundary condition, describe Quasi-Birth-Death processes also known as QBD. The more general class of stochastic matrices obtained by replacing the entries of (2) with matrix blocks describes *Non-Skip-Free* Markov chains [20]. By means of the reblocking technique the latter apparently wider class can be reduced to an M/G/1 with blocks of larger size.

We refer the reader to the books [37], [38], [33] for a more detailed treatment of this kind of stochastic processes and to the book [5] for a systematic treatment of the algorithmic and computational aspects of these problems. Here we point out that M/G/1/, G/M/1 and QBD stochastic processes model a very wide variety of problems ranging from telecommunications, computer networks, mobile telephones, decision processes, medical sciences, risk analysis, etc.

For a more detailed description on recent applications, methodologies, design and analysis of algorithms we refer the reader to the volume [4] with the proceedings of the conference "Matrix Analytic Methods for Stochastic Processes", Pisa, June 22–26 2005.

2. Computational problems

Solving a Markov chain, i.e., computing the vector π seems, at first glance, a trivial problem at least for finite matrices. In fact, the condition $\pi^T P = \pi^T$ can be rewritten in the form of a homogeneous singular linear system $(I - P)^T \pi = 0$ which can be solved by means of general linear algebra tools as the LU factorization of $I - P^T$ [46, 23]. The theory of M-matrices [47], [3] guarantees the existence of the LU factorization, moreover, the computation of L and U such that $(I - P) = LU$ by means of Gaussian elimination becomes a numerically stable computation if complemented with the GTH trick for removing possible cancellation [24].

Once the factor U has been computed, the triangular system $U\pi = 0$ can be easily solved by means of back substitution. The overall cost is dominated by Gaussian elimination and amounts to $\frac{2}{3}n^3 + O(n^2)$ arithmetic operations.

Actually, Gaussian elimination is a very effective tool if applied to moderately large matrices. If one tries to apply this method to the page rank problem some (expected) difficulty is encountered. For $n = 8.5 \cdot 10^9$ the overall number of operations required by Gaussian elimination is roughly 4.1×10^{29}. Let us try to estimate the CPU time needed to solve the page rank problem with Gaussian elimination by using a fast computer.

The fastest computer available nowadays is the Blue Gene by IBM. It has a speed of 360 teraflops, that is, it can execute 3.6×10^{14} arithmetic operations per second. In order to execute 4.1×10^{29} arithmetic operations it would need about 1.1×10^{15} seconds corresponding to more than 36 million years. Indeed, Google cannot rely on Gaussian elimination in order to sort the web pages according to their rank.

In the case where the Markov chain is infinite the computation is even more complicated and different techniques based on the peculiarity of the problem must be used in order to design highly efficient solution algorithms.

In the next sections we will deal with infinite problems encountered in queueing models and show the main ideas and techniques used for their solution. Concerning the methods designed for solving the page rank problem we refer the reader to the recent papers [1, 14, 29, 36, 40, 44]. Here we just want to point out that this problem is currently receiving a great interest from researchers working in numerical linear algebra for the importance of its applications and for the interesting numeric tools that are used for its solution.

A crucial property is that the connectivity matrix H is a highly sparse matrix. In fact, the number of nonzero elements in each line, i.e., the number of links which are listed in a web page, is greatly inferior to the number n of pages available in the web. Therefore, multiplying the matrix H as well as the matrix P by a vector has a cost which is bounded from above by a small multiple of n so that the matrix vector multiplication can be performed in less than a second of CPU time even with a PC despite the huge size of the matrix. This is the basis for the design of efficient algorithms for approximating π in few hours of CPU time based on the elementary operations of matrix-vector multiplication.
The semi-infinite matrices encountered in queuing models have the block Toeplitz block Hessenberg structure. Exploiting the specific properties of this structure is the key idea for designing very efficient solution algorithms.

In general, in order to solve large scale scientific problems it is not convenient to use general methods. As in the case of the page rank problem, the large CPU execution time makes the problem practically unsolvable. Only through the analysis of the peculiarity of the problem, which can be better understood in terms of mathematical (matrix) structures, it is possible to design effective solution algorithms despite the large size of the problem. The analysis of mathematical structures is not only the way for designing highly efficient algorithms, but is also the main source of enjoyment and motivation of the research in mathematics. In this regard, it is worth citing the following comment by Alexander Grothendieck, Fields medal in 1966:

If there is one thing in mathematics that fascinates me more than anything else (and doubtless always has), it is neither number nor size, but always form. And among the thousand-and-one faces whereby form chooses to reveal itself to us, the one that fascinates me more than any other and continues to fascinate me, is the structure hidden in mathematical things.

3. M/G/1 Markov chains: Wiener-Hopf factorizations and the Ramaswami formula

The aim of this section is to give the flavor of the mathematical tools used in solving M/G/1 stochastic processes without delving into technical details. Throughout the section we follow [5] to which we refer the readers who wish to have more information on this topic.

Let us recall that our main goal is to compute a finite number of components of the infinite vector $\pi = (\pi_i) \in \mathbb{R}^{\mathbb{N}}$ such that

$$\pi^T(I - P) = 0, \quad \pi_i \geq 0, \quad \sum_{i \in \mathbb{N}} \pi_i = 1, \tag{4}$$

where, assuming more general boundary conditions, we set

$$P = \begin{bmatrix} \widehat{A}_0 & \widehat{A}_1 & \widehat{A}_2 & \dots & \dots \\ \widehat{A}_{-1} & A_0 & A_1 & A_2 & \ddots \\ & A_{-1} & A_0 & A_1 & \ddots \\ O & & \ddots & \ddots & \ddots \end{bmatrix}, \tag{5}$$

where the blocks are $m \times m$ matrices.

3.1. The Ramaswami formula

A classical tool for solving the infinite system (4) is the Ramaswami formula which has been proved in [38] by using probabilistic arguments. Here we describe it by relying only on elementary linear algebra tools.

Denote by $\widetilde{P}$ the submatrix of P obtained by removing the first block row and the first block column of P. The matrix $\widetilde{P}$ is purely block Toeplitz and in upper Hessenberg form. Assume that the following UL factorization holds

$$I - \widetilde{P} = \begin{bmatrix} U_0 & U_1 & U_2 & \dots \\ & U_0 & U_1 & \ddots \\ & & \ddots & \ddots \\ O & & & \ddots \end{bmatrix} \begin{bmatrix} I & & & O \\ -G & I & & \\ & -G & I & \\ O & & \ddots & \ddots \end{bmatrix} = UL \tag{6}$$

where the left factor U is a block upper triangular block Toeplitz matrix and the right factor L is a block lower triangular block Toeplitz matrix having only two nonzero block diagonals.

The system (4), rewritten as

$$\left[\begin{array}{cc} \widehat{\pi}^T & \widetilde{\pi}^T \end{array}\right]\left[\begin{array}{c|c} I-\widehat{A}_0 & -\widehat{A}_1 \quad -\widehat{A}_2 \quad \ldots \\ \hline -\widehat{A}_{-1} & \\ 0 & UL \\ \vdots & \end{array}\right]=0$$

for $\widehat{\pi} \in \mathbb{R}^m$, $\widetilde{\pi} \in \mathbb{R}^{\mathbb{N}}$, yields

$$\left\{\begin{array}{l} \widetilde{\pi}^T UL = \widehat{\pi}^T \left[\begin{array}{ccc} \widehat{A}_1 & \widehat{A}_2 & \ldots \end{array}\right] \\ \widehat{\pi}^T (I-\widehat{A}_0 - \left[\begin{array}{ccc} \widehat{A}_1 & \widehat{A}_2 & \ldots \end{array}\right] L^{-1}U^{-1}\left[\begin{array}{c} \widehat{A}_{-1} \\ 0 \\ \vdots \end{array}\right]) = 0 \end{array}\right.$$

so that it can be reduced to the finite system

$$\widehat{\pi}^T\left[I-\widehat{A}_0-\left(\sum_{i=1}^{\infty}\widehat{A}_i G^{i-1}\right)U_0^{-1}\widehat{A}_{-1}\right]=0 \tag{7}$$

and to the two block triangular (infinite) block Toeplitz systems

$$\widetilde{\pi}^T U = \mathbf{c}^T, \quad \mathbf{c}^T L = \widehat{\pi}^T[\widehat{A}_1\ \widehat{A}_2\ \ldots]. \tag{8}$$

Let $||\cdot||$ be any matrix norm. If $||G^i||$ is uniformly bounded from above for any $i \in \mathbb{N}$ then the above systems are solvable and the sequence $||U_i||$ decays to zero [5]. This makes easily solvable the infinite triangular systems (8).

Under weak assumptions on the Markov chain, such a matrix G exists and can be computed as the solution of the matrix equation

$$X=\sum_{i=-1}^{+\infty}X^{i+1}A_i \tag{9}$$

of minimal spectral radius [5]. Partitioning the vector π into subvectors π_i of size m for $i=0,1,\ldots$, we find that $\widehat{\pi}=\pi_0$ and the solution of (7) and (8) provides the equations

$$\left\{\begin{array}{l} \pi_0^T(I-\widehat{A}_0-A_1^*U_0^{-1}\widehat{A}_{-1})=0 \\ \pi_1^T=\pi_0^T A_1^* U_0^{-1} \\ \pi_i^T=(\pi_0^T A_i^*-\sum_{j=1}^{i-1}\pi_j^T U_{i-j})U_0^{-1}, \quad i\geq 2 \\ A_i^*=\sum_{j=0}^{+\infty}\widehat{A}_{i+j}G^i \end{array}\right. \tag{10}$$

known as Ramaswami formula. By means of these expressions it is possible to compute as many components of π as needed. The complexity of the straightforward application of (10) can be substantially decreased by using the Toeplitz matrix technology with the FFT [34].

Sometimes changing the point of view helps to better understand the problems of the world. If we look at the problem from a functional point of view we have more insights of the Ramaswami formula; this enables one to design faster and numerically stable algorithms for its computation.

3.2. Wiener-Hopf factorizations

Let us associate with the matrix $I - \widetilde{P}$ the matrix Laurent series

$$S(z) = I - \sum_{i=-1}^{\infty} z^i A_i$$

Since $A_i \geq 0$ and $\sum_{i=-1}^{+\infty} A_i$ is finite, then $S(z)$ belongs to the *Wiener algebra* [21]

$$\mathcal{W} = \{F(z) = \sum_{i=-\infty}^{+\infty} z^i F_i : \quad F_i \in \mathbb{C}^{m\times m}, \quad \sum_{i=-\infty}^{+\infty} ||F_i|| < +\infty\}.$$

Matrix valued functions $F(z)$ which are analytic in the annulus

$$\mathcal{A} = \{z \in \mathbb{C} : \quad r < |z| < R\}, \quad r < 1 < R \tag{11}$$

clearly belong to the Wiener algebra and have an exponential decay to zero of the norm $||F_i||$ of their coefficients for $i \to \pm\infty$.

The following definition plays an important role [13], [21], [5].

DEFINITION 1. A factorization of $F(z) \in \mathcal{W}$ of the kind

$$F(z) = U(z)L(z), \quad U(z) = \sum_{i=0}^{+\infty} z^i U_i, \quad L(z) = \sum_{i=0}^{+\infty} z^{-i} L_{-i}, \quad |z| = 1,$$

where $U(z), L(z) \in \mathcal{W}$, $\det U(z), \det L(z^{-1}) \neq 0$ for $|z| \leq 1$, is called *canonical (Wiener-Hopf) factorization*. The factorization is *weak* if $U(z)$ and $L(z^{-1})$ are non-singular for $|z| < 1$, and there exists some z of modulus 1 such that $\det U(z) = 0$ or $\det L(z) = 0$.

Non-canonical factorizations have the form $U(z)D(z)L(z)$ where $D(z) = \mathrm{diag}(z^{\kappa_1}, \dots, z^{\kappa_m})$, and the integers κ_i are called *partial indices*. A canonical factorization is a Wiener-Hopf factorization with null partial indices.

Observe that if $F(z)$ is analytic on the annulus $\mathcal{A}$, the exponential decay of its coefficients implies the exponential decay to zero of the coefficients of $U(z)$ and $L(z)$. Moreover, a canonical factorization provides a splitting of the zeros of $\det F(z)$ with respect to the unit circle. In fact, $\det F(z) = \det U(z) \det L(z)$ and $\det U(z)$ has zeros inside the unit disk, $\det L(z)$ has zeros outside the unit disk. In the case where $F(z)$ is a Laurent polynomial, the canonical Wiener Hopf factorization of $F(z)$ is obtained by splitting $F(z)$ as a product of a polynomial in z with zeros of modulus less than 1 and a polynomial in z^{-1} with zeros of modulus greater than 1.

A canonical factorization $F(z) = U(z)L(z)$ can be rewritten in matrix form as a UL factorization of semi-infinite block Toeplitz matrices:

$$\begin{bmatrix} F_0 & F_1 & F_2 & F_3 & \dots \\ F_{-1} & F_0 & F_1 & F_2 & \ddots \\ F_{-2} & F_{-1} & F_0 & F_1 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} = UL =$$

$$\begin{bmatrix} U_0 & U_1 & U_2 & \dots \\ & U_0 & U_1 & \ddots \\ O & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} L_0 & & & O \\ L_{-1} & L_0 & & \\ L_{-2} & L_{-1} & L_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Observe that the Ramaswami formula (10) which relies on the UL factorization (6) of the matrix $I - \widetilde{P}$ is ultimately reduced to computing a weak canonical factorization of $S(z) = I - \sum_{i=-1}^{+\infty} z^i A_i$. In fact, from the stochasticity of $A = \sum_{i=-1}^{+\infty} A_i$ one has $A\mathbf{e} = \mathbf{e}$, which yields

$$S(1)\mathbf{e} = (I - \sum_{i=-1}^{+\infty} A_i)\mathbf{e} = 0,$$

that is, the matrix $S(z)$ is singular for $z = 1$ and the factorization (6) leads to a weak canonical factorization of $S(z)$. Under very mild conditions on the queueing problem one has that $z = 1$ is the only singular point of modulus 1 for the matrix $S(z)$ [5]. Without loss of generality we may assume this condition valid throughout the paper.

Wiener-Hopf factorizations always exist provided that $F(z) \in \mathcal{W}$ is such that $\det F(z) \neq 0$ for $|z| = 1$. But canonical factorizations not always exist. Condition for their existence are widely studied and there is a consolidated literature on this subject, we refer the reader to [21] for a general treatment. More recently, the interest has been addressed to the design of algorithms for computing canonical factorizations [5]. A part which is less investigated concerns weak factorizations.

From the computational point of view the main effort in solving an M/G/1 Markov chain is the computation of a weak canonical factorization. This motivates the interest on the design of algorithms for its computation.

In the next section we report the main results concerning the computational issues of (weak) canonical factorizations, for more details we refer the reader to the book [5].

4. Algorithmic approaches

There are two main approaches for computing canonical factorizations of matrix function of the kind $F(z) = \sum_{i=-N}^{+\infty} z^i F_i$ for a positive integer N. A first approach reduces the problem to solving a matrix equation, a second approach consists in inverting a matrix Laurent series and solving a finite block Toeplitz system [5].

Without loss of generality we may assume $N = 1$. Otherwise, by applying the reblocking technique, we may reduce (2) to (3) with an increasing of the block size by a factor of N.

For notational consistency with the problems in Markov chains we assume that

$$S(z) = I - \sum_{i=-1}^{\infty} z^i A_i. \tag{12}$$

THEOREM 1. *If the matrix valued function $S(z) = I - \sum_{i=-1}^{\infty} z^i A_i \in \mathcal{W}$ has the canonical factorization*

$$S(z) = U(z)(I - z^{-1}G) \tag{13}$$

then G is the unique solution of the the matrix equation

$$X = \sum_{i=-1}^{\infty} A_i X^{i+1} \tag{14}$$

of minimum spectral radius $\rho(G) < 1$. Conversely, if there exists a solution G of (14) *such that $\rho(G) < 1$ and the power series* $\det(zS(z))$ *has exactly m roots of modulus less than 1 and is nonzero for $|z| = 1$, then $S(z)$ has the canonical factorization* (13).

The solution G of (14) with minimal spectral radius is called *minimal solution*.

A similar result [5] relates the weak canonical factorization of $S(z)$ and the existence of a minimal solution G such that $\rho(G) = 1$ and $||G^k||$ is uniformly bounded, provided that both $S(z)$ and its first derivative $S'(z)$ belong to the Wiener algebra. The latter condition clearly holds if $S(z)$ is analytic in the annulus $\mathcal{A}$ of (11).

The following result relates a canonical factorization to the coefficients of a suitable matrix Laurent series [5, 7].

THEOREM 2. *Let $F(z) = \sum_{i=-N}^{+\infty} z^i F_i \in \mathcal{W}$, $N \geq 1$, assume that there exists a canonical factorization*

$$F(z) = U(z)L(z),$$

where without loss of generality we may assume $L_0 = I$. Then $H(z) = F(z)^{-1} \in \mathcal{W}$, moreover if $H(z) = \sum_{i=-\infty}^{+\infty} z^i H_i$, $T_q = (H_{j-i})_{i,j=1,q}$, $q > N$, one has

$$[L_0\ 0\ \dots] = U_0[L_{-q+1}\ \dots\ L_0]T_q.$$

From the above theorem it follows that computing a canonical factorization can be reduced to

1. computing the central coefficients of the matrix Laurent series $H(z) = F^{-1}$;

2. solving a $q \times q$ finite block Toeplitz system for $q > N$.

If the matrix Laurent series $F(z)$ is analytic and nonsingular for $z \in \mathcal{A}$, then also $H(z) = F(z)^{-1}$ is analytic for $z \in \mathcal{A}$. The analyticity of $H(z)$ implies the exponential decay of H_i for $i \to \pm\infty$. Numerically, one may look at $H(z)$ as to a matrix Laurent polynomial of sufficiently large degree. Therefore, its coefficients can be approximated by means of an evaluation/interpolation technique described in the following algorithm.

Algorithm: Evaluation/interpolation

1. Choose an integer $M > 0$;

2. evaluate $F(z)$ at the Mth roots of 1 and obtain matrices $W_i = F(\omega_i)$, $i = 1 : M$;

3. compute $V_i = W_i^{-1}$, $i = 1 : M$

4. interpolate to the pairs (ω_i, V_i) and get approximations to the matrix coefficients H_i, $i = 1 : M$;

5. if the approximations are satisfactory, then output H_i, $i = 1 : M$ and stop; otherwise set $M = 2M$ and repeat from step 2.

Observe that this approach cannot be applied if $\det F(z) = 0$ for some z of modulus 1. Unfortunately, this is the case of problems coming from Markov chains since for $F(z) = S(z)$ one has $\det S(1) = 0$. In the next section it is shown how this drawback can be overcome.

5. Removing singularities

We describe a suitable transformation for turning a weak canonical factorization of the function $S(z)$ of (12) such that $S(1)\mathbf{e} = 0$ into a canonical factorization. The idea, originally introduced in [28] but already contained in different form in the previous paper [17] has been developed and perfected in [8, 12, 5]. it consists in constructing a new matrix valued function $\widetilde{S}(z) \in \mathcal{W}$ which is nonsingular for $z = 1$ and such that $\det \widetilde{S}(z) = 0$ if $\det S(z) = 0$, $z \neq 1$.

Assume that $z = 1$ is the only value of modulus 1 such that $\det S(z) = 0$ and assume that this root is also simple. Let $\mathbf{u}$ be any vector such that $\mathbf{u}^T\mathbf{e} = 1$ and set

$$\widetilde{S}(z) = S(z)(I - z^{-1}Q), \quad Q = \mathbf{e}\mathbf{u}^T.$$

In [5] it is shown that if $S(z), S'(z) \in \mathcal{W}$ then also $\widetilde{S}(z) \in \mathcal{W}$, moreover, $\det \widetilde{S}(z) \neq 0$ for $|z| = 1$ and $\widetilde{S}(z) = I - \sum_{i=-1}^{+\infty} z^i \widetilde{A}_i$ for suitable $m \times m$ blocks $\widetilde{A}_i$. Moreover, if

$S(z) = U(z)L(z)$ is a weak canonical factorization of $S(z)$, then $\widetilde{S}(z) = \widetilde{U}(z)\widetilde{L}(z)$ is a canonical factorization of $\widetilde{S}(z)$ where $\widetilde{U}(z) = U(z)$ and $\widetilde{L}(z) = L(z)(I - z^{-1}Q)$.

With this technique the computation of a weak canonical factorization of $S(z)$ is reduced to computing a canonical factorization of $\widetilde{S}(z)$. The latter computation can be carried out by means of the evaluation/interpolation technique complemented by Theorem 2.

6. Solving a matrix equation: traditional algorithms

The most direct way for solving the matrix equation (14) is using fixed point iterations. A very natural way is to consider sequences generated by

$$X^{(k+1)} = \sum_{i=-1}^{+\infty} A_i (X^{(k)})^{i+1} \tag{15}$$

given an initial approximation $X^{(0)}$. Other iterations can be generated with similar formulas like

$$X^{(k+1)} = (I - A_0)^{-1}(A_{-1} + \sum_{i=1}^{+\infty} A_i (X^{(k)})^{i+1}) \tag{16}$$

or like

$$X^{(k+1)} = (I - \sum_{i=0}^{+\infty} A_i (X^{(k)})^{i})^{-1} A_{-1}. \tag{17}$$

For the sequences generated by all the above iterations, convergence is linear as long as the Markov chain is positive recurrent, it is sublinear if the Markov chain is null recurrent. Choosing $X^{(0)} = 0$ provides a monotonic convergence to the minimal solution G. Choosing $X^{(0)}$ stochastic, say $X^{(0)} = I$, provides a higher convergence rate, moreover the iteration (17) is faster than (16) which is faster than (15) [35]. For more information, we refer the reader to the books [37, 38, 5], and to the papers [42, 31], a very accurate analysis for Non-Skip-Free problems is carried out in the paper by Gail, Hantler and Taylor [20], a systematic and ultimate treatment is performed in [35].

The main drawback of fixed-point iterations is the excessive slowdown that is encountered by these methods for problems which are close to singularity. In fact, for positive recurrent Markov chains having a “drift” close to zero the convergence slows down and the number of iteration is so large that the CPU time does not allow one to provide a solution in real time.

This drawback is overcome by using quadratically convergent algorithms which have been designed starting from the ’90s, and rely on the divide and conquer technique. In [32], Latouche and Ramaswami introduce the algorithm of Logarithmic Reduction (LR) for QBDs, which is an adjustment of the Cyclic Reduction (CR) algorithm introduced in the early ’70s by Buzbee, Golub and Nielson [15] for solving elliptic boundary value problems. In [30] Li and Sheng introduce the folding algorithm for QBDs which is essentially the same as LR. In [9] the Cyclic Reduction algorithm is revisited and adapted for solving more general M/G/1 problems. A systematic analysis and great improvements of CR are carried out in the papers [6, 10, 28, 7, 8, 25, 26, 27]; less efficient divide and conquer techniques are introduced in [45].

The very nice feature of this class of algorithm is the quadratic convergence to the minimal solution for positive recurrent Markov chains. Convergence is still valid for null recurrent Markov chains; in this case, even though the convergence is linear, the "shift" strategies of Section 5 can be applied for making it quadratic [5, 12]. The basic concepts of CR are introduced in the next section.

7. Solving a matrix equation: advanced algorithms

Formally, the *nonlinear* matrix equation

$$-A_{-1} + (I - A_0)X - A_1X^2 - A_2X^3 - \cdots = 0$$

can be rewritten as a *linear* system defined by a block Toeplitz matrix in block Hessenberg form:

$$\begin{bmatrix} I - A_0 & -A_1 & -A_2 & -A_3 & \dots \\ -A_{-1} & I - A_0 & -A_1 & -A_2 & \ddots \\ 0 & -A_{-1} & I - A_0 & -A_1 & \ddots \\ & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^2 \\ X^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{-1} \\ 0 \\ 0 \\ \vdots \end{bmatrix}.$$

The tax that we have to pay for this linearization is the infinite size of the linear system. Observe that any algorithm for solving the above system provides a solver for the matrix equation (9).

For the sake of simplicity let us consider the case of a QBD where the system turns into

$$\begin{bmatrix} D & C & & O \\ B & D & C & \\ O & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^2 \\ X^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} -B \\ 0 \\ \vdots \end{bmatrix} \tag{18}$$

with $B = -A_{-1}$, $D = I - A_0$, $C = -A_1$.

The Cyclic Reduction algorithm is a natural and powerful method for the above system. It was introduced in [15] for solving finite system discretizing partial differential equations of elliptic type. Its effectiveness still holds for infinite systems. By means of an even/odd permutation of block rows and columns, the system (18) is transformed into

$$\left[\begin{array}{ccc|ccc} D & & & B & C & \\ & D & & & B & \ddots \\ & & \ddots & & \ddots & \ddots \\ \hline C & & & D & & \\ B & C & & & D & \\ & \ddots & \ddots & & & \ddots \end{array}\right]\left[\begin{array}{c} X^2 \\ X^4 \\ \vdots \\ \hline X \\ X^3 \\ \vdots \end{array}\right]$$

One step of block Gaussian elimination leads to the system

$$\begin{bmatrix} \widehat{D}_1 & C_1 & & O \\ B_1 & D_1 & C_1 & \\ O & \ddots & \ddots & \ddots \end{bmatrix}\begin{bmatrix} X \\ X^3 \\ X^5 \\ \vdots \end{bmatrix} = \begin{bmatrix} -B \\ 0 \\ \vdots \end{bmatrix}$$

where $D_1 = D - BD^{-1}C - CD^{-1}B$, $B_1 = -BD^{-1}B$, $C_1 = -CD^{-1}C$, $\widehat{D}_1 = D - CD^{-1}B$.

Surprisingly, the original structure of the system is almost maintained. The same procedure can be recursively applied to the new block tridiagonal system. This provides the following sequence of block tridiagonal Toeplitz-like systems

$$\begin{bmatrix} \widehat{D}_k & C_k & & O \\ B_k & D_k & C_k & \\ O & \ddots & \ddots & \ddots \end{bmatrix}\begin{bmatrix} X \\ X^{2^k+1} \\ X^{2\cdot 2^k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B \\ 0 \\ \vdots \end{bmatrix}$$

where

$$\begin{aligned} D_{k+1} &= D_k - B_k D_k^{-1} C_k - C_k D_k^{-1} B_k, & \widehat{D}_{k+1} &= D_k - C_k D_k^{-1} B_k, \\ B_{k+1} &= -B_k D_k^{-1} B_k, & C_{k+1} &= -C_k D_k^{-1} C_k. \end{aligned}$$

In the finite case, the size of the system is halved at each recursive step of the algorithm. In this way, once a 1×1 system is obtained after a finite number of recursive steps, the back substitution technique allows one to compute all the components of the solution. In the infinite case we obtain an infinite sequence of infinite systems so that the back-substitution stage in principle can never be applied. However, we may prove nice properties of this matrix sequence which enable one to compute the solution G in a very effective way.

A first observation is that the unknown X of our matrix equation $CX^2 + DX + B = 0$ occupies the first position of the unknown vector of the block tridiagonal system at all the recursive stages. As a consequence, the first equation of the kth system of the sequence provides the expression

$$X = -\widehat{D}_k^{-1}(B + C_k X^{2^k+1}) \tag{19}$$

provided that $\widehat{D}_k$ is nonsingular.

A second and crucial property is that, under weak conditions the matrix C_k converges to zero double exponentially, moreover $||X^{2^k+1}||$ is uniformly bounded from above by a constant and $\widehat{D}_k^{-1}$ has a finite limit.
This is the basis to prove that $\lim_k -\widehat{D}_k^{-1}B = G$ and the convergence is quadratic. Thus the expression $-\widehat{D}_k^{-1}B$ provides an efficient approximation to the minimal solution G. We refer the reader to the book [5] for a systematic treatment of these convergence properties.

A better understanding of CR is provided by the functional formulation of this method.

7.1. Functional formulation of CR

Let us change once again our point of view and associate with the sequence of triples (C_k, D_k, B_k) generated by CR the sequence of Laurent matrix polynomials

$$\varphi_k(z) = z^{-1}B_k + D_k + zC_k$$

with $\varphi_0(z) = S(z) = -z^{-1}A_{-1} + I - A_0 - zA_1$. Then one can easily prove that

$$\varphi_{k+1}(z^2) = -\varphi_k(z)D_k^{-1}\varphi_k(-z)$$

It is surprising to find out that the above functional relation is nothing else but the Graeffe iteration [39] adjusted to matrix Laurent polynomials with three terms.

Observe also that looking at the matrix inverse $\psi_k(z) = \varphi_k(z)^{-1}$, one deduces the equation

$$\psi_{k+1}(z^2) = \frac{1}{2}(\psi_k(z) + \psi_k(-z)) \tag{20}$$

which makes substantially change the perspective. In fact, denoting $\psi_0(z) = \phi_0(z)^{-1} = \sum_{i=-\infty}^{+\infty} z^i H_i$, then from (20) one finds that

$$\psi_k(z) = \sum_{i=-\infty}^{+\infty} z^i H_{i\cdot 2^k}.$$

The analyticity of $\psi_0(z)$ implies the exponential decay of the coefficients of $\psi_0(z)$ and consequently, the double exponential decay to zero of the coefficients of $\psi_k(z)$ with respect to k. This is the key property at the basis of convergence of CR. Moreover, the wider is the width of the domain of analyticity $\mathcal{A}$ of $\psi_0(z)$ the faster is the convergence of the method.

Conditions of applicability of CR can be easily proved. In fact, break-down is possible only if $\det D_k = 0$ for some k. The functional formulation allows one to state very mild conditions under which the nonsingularity is ensured [12].

The functional formulation allows one to generalize CR to the case of M/G/1 Markov chains where the matrix equation has the more general form (14).

7.2. Extension to the M/G/1 case

Consider the equation (14) and the matrix Laurent series

$$\varphi_0(z) = -z^{-1}A_{-1} + (I - A_0) - zA_1 - z^2A_2 - \cdots$$

Define recursively the following sequence of matrix Laurent series

$$\varphi_{k+1}(z^2) = -\varphi_k(z)\,[\varphi_k(z)]_{even}^{-1}\,\varphi_k(-z) \tag{21}$$

with

$$[\varphi_k(z)]_{even} = \tfrac{1}{2}[\varphi_k(z) + \varphi_k(-z)]$$
$$[\varphi_k(z)]_{odd} = \tfrac{1}{2z}[\varphi_k(z) - \varphi_k(-z)]$$

where we assume that $[\varphi_k(z)]_{even}$ is invertible for $|z| = 1$ for any k. Define also the auxiliary sequence of matrix Laurent series

$$\widehat{A}^{(k+1)}(z^2) = [\widehat{A}^{(k)}(z)]_{odd} + [\widehat{A}^{(k)}(z)]_{even}[\varphi_k(z)]_{even}^{-1}(I - [\varphi^{(k)}(z)]_{odd}) \tag{22}$$

with $\widehat{A}^{(0)}(z) = \sum_{i=0}^{+\infty} z^{i-1}A_i$. Then,

$$X = (I - \widehat{A}_0^{(k)})^{-1}(A_{-1} + \sum_{i=1}^{+\infty} \widehat{A}_i^{(k)} X^{i\cdot 2^k+1})$$

where $\widehat{A}^{(k)}(z) = \sum_{i=-1}^{+\infty} z^i \widehat{A}_i^{(k)}$.
Under mild conditions generally satisfied in the applications, one may prove that the sequence is well defined, the summation $\sum_{i=1}^{+\infty} \widehat{A}_i^{(k)} X^{i\cdot 2^k+1}$ converges quadratically to zero, the matrix sequence $(I - \widehat{A}_0^{(k)})^{-1}$ is bounded and quadratically convergent. This implies that

$$\lim_k (I - \widehat{A}_0^{(k)})^{-1}A_{-1} = G$$

and that convergence is quadratic.

This property is the basis for extending the algorithm of CR to matrix equations of the kind (14) defined by means of a matrix Laurent series and provides an effective tool for their solution. In matrix form, CR provides also a powerful technique for solving block Toeplitz systems in block Hessenberg form. Once again, we refer the reader to the book [5] for more information.

7.3. Computational aspects of CR

The convergence of CR depends on the width r/R of the annulus $\mathcal{A}$ of (11). In [5] it is proved that the approximation G_k to the solution G obtained after k steps of CR is such that $||G_k - G|| = O((r/R)^{2^k})$. This shows that the wider is the width of $\mathcal{A}$ the faster is the convergence. Since the values of r and R are given by the moduli of the largest zero of $\det \varphi_0(z)$ in the unit disk, and the smallest zero outside the unit disk, respectively, for positive recurrent Markov chains we have $r = 1$ and $R > 1$. The shift technique of Section 5 allows one to remove the zero $r = 1$ and replace it with the modulus of the largest zero of modulus strictly less than 1. For null recurrent Markov chains where $r = R = 1$ the annulus $\mathcal{A}$ is the empty set and convergence of CR turns to linear. However, applying once again the shift technique of Section 5 we obtain a new matrix equation where the analyticity domain of the function is a nonempty annulus so that convergence is still quadratic.

Concerning the computational cost of CR, it is interesting to observe that for a quadratic matrix equation, which model a QBD process, one step of CR costs just one $m \times m$ matrix inversion and six $m \times m$ matrix multiplications. For a general M/G/1 problem the cost depends on the decay to zero of the coefficients of $\varphi_k(z)$. In fact, the equations (21) and (22) can be implemented by means of the evaluation/interpolation technique where the number of interpolation knots is depending on the *numerical degree* of the matrix Laurent series which in turn depends on the decay speed of the coefficients. From the practice of computations it results that the *point-wise* implementation of CR of [10] is much more effective than the implementation based on the matrix formulation [6].

8. Conclusions and open problems

We have shown that the analysis of structured matrices together with a blended combination of different tools from linear algebra, complex analysis and operator theory allows one to design effective algorithms for solving complex Markov chains. The lion role is played by the canonical Wiener-Hopf factorization of matrix valued functions of a complex variable. The computation of the latter factorization is reduced to solving a matrix equation or to inverting a matrix Laurent power series. A known tool from numerical linear algebra, the cyclic reduction algorithm, has been used and interpreted in a functional form. This interpretation has led to its generalization to M/G/1 problems. A shift technique has been described for removing a singularity and for reducing a weak canonical factorization to a canonical factorization. This technique enables one to widen the domain of analyticity of a matrix function improving in this way the convergence speed of cyclic reduction.

There are several open problems which deserve some attention. In certain queueing models functions of the kind $A(z) = \sum_{i=-\infty}^{+\infty} z^i A_i$ are encountered. Efficient algorithms for dealing with this kind of functions are not known.

Another interesting situation is the case where the size of the blocks A_i is infinite. Attempts to deal with this case of “operator valued functions” have been tried in [8, 11], but much has to be done especially from the algorithmic point of view. Blocks of infinite size are encountered in certain models, moreover, in [43] it is proved that an M/G/1 problem with finite blocks can be reduced to a QBD with infinite blocks.

The theory of (canonical) Wiener-Hopf factorization is well understood. There is a wide literature on matrix polynomials [22] and matrix power series with many theoretical results concerning Wiener-Hopf factorizations [21]. On the other hand very few results concerning weak factorizations are available [5], [17] and much has still to be done.

Another interesting theoretical issue concerns the applicability of cyclic reduction. Sufficient conditions for the applicability of CR are known [12] but a complete characterization is still missing.

Great interest has been recently addressed to the problem of solving fluid queues where the set of states is continuous [41], [2]. Certain Riccati equations are encountered in the solution of these models and their reduction to unilateral quadratic matrix equations seems to be an effective approach where quadratically convergent algorithms like CR can be applied.

References

[1] ARASU A., NOVAK J., TOMKINS A. AND TOMLIN J. *Page Rank computation and the structure of the Web*, `http://www2002.org/CDROM/poster/173.pdf`, 2002.

[2] BEAN N.G., O'REILLY M.M. AND TAYLOR P.G., *Algorithms for return probabilities for stochastic fluid flows*, Stochastic Models **21** (2005), 149–184.

[3] BERMAN A. AND PLEMMONS R.J., *Nonnegative matrices in the mathematical sciences*, Classics in Applied Mathematics **9**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA 1994.

[4] BINI D.A., LATOUCHE G. AND MEINI B. *Fifth International Conference on Matrix Analytic Methods* Stochastic Models **21**, (2005).

[5] BINI D.A., LATOUCHE G. AND MEINI B. *Numerical Methods for Structured Markov Chains*, Oxford University Press, 2005.

[6] BINI D.A. AND MEINI B. *On the solution of a nonlinear matrix equation arising in queueing problems*, SIAM J. Matrix Anal. Appl. **17** (1996), 906–926.

[7] BINI D.A., GEMIGNANI L. AND MEINI B. *Computations with infinite Toeplitz matrices and polynomials*, Linear Algebra Appl. **343/344** (2002), 21–61.

[8] BINI D.A., GEMIGNANI L. AND MEINI B. *Solving certain matrix equations by means of Toeplitz computations: algorithms and applications*, Contemporary. Mathematics **323** (2003), Amer. Math. Soc., Providence, RI, 151–167.

[9] BINI D.A. AND MEINI B. *On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems*, in: "Computations with Markov Chains" (W. J. Stewart ed.), Kluwer Academic Publishers, Boston 1995, 21–38.

[10] BINI D.A. AND MEINI B. *Improved cyclic reduction for solving queueing problems*, Numer. Algorithms, **15** (1997) 57–74.

[11] BINI D.A., MEINI B. AND RAMASWAMI V., *Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G*, in: "Advances in Algorithmic Methods for Stochastic Models, Proceedings of the Third Conference on Matrix Analytic Methods", (G. Latouche and P. Taylor Eds.), Notable Publications, New Jersey 2000.

[12] BINI D.A., MEINI B. AND SPITKOWSKY I.M., *Shift techniques and canonical factorizations in the solution of M/G/1-type Markov chains*, Stochastic Models **21** (2005),279–302.

[13] BÖTTCHER A. AND SILBERMANN B. *Introduction to large truncated Toeplitz matrices*. Universitext. Springer-Verlag, New York, 1999.

[14] BREZINSKI C., REDIVO-ZAGLIA M. AND SERRA-CAPIZZANO S., *Extrapolation methods for PageRank computations*, C. R. Math. Acad. Sci. Paris **340** (2005), 393–397.

[15] BUZBEE B.L., GOLUB G.H. AND NIELSON C.W., *On direct methods for solving Poisson's equations*, SIAM J. Numer. Anal. **7** (1970), 627–656.

[16] ÇINLAR E., *Introduction to stochastic processes*, Prentice-Hall, Englewood Cliffs, NJ 1975.

[17] DUKHOVNY A., *Complex analytic factorization and partial indices of matrix functions associated with Markov controlled random walks*, Commun. Statist. Stochastic Models **12** (1996), 683–697.

[18] GAIL H.R., HANTLER S.L. AND TAYLOR B.A., *Solutions of the basic matrix equation for $M/G/1$ and $G/M/1$ type Markov chains*, Comm. Statist. Stochastic Models **10** (1994), 1–43.

[19] GAIL H.R., HANTLER S.L. AND TAYLOR B.A., *Spectral analysis of $M/G/1$ and $G/M/1$ type Markov chains*, Adv. in Appl. Probab. **28** (1996), 114–165.

[20] GAIL H.R., HANTLER S.L. AND TAYLOR B.A., *Non-skip-free $M/G/1$ and $G/M/1$ type Markov chains*, Adv. in Appl. Probab. **29** (1997), 733–758.

[21] GOHBERG I., KAASHOEK M.A. AND SPITKOVSKY I.M., *An overview of matrix factorization theory and operator applications*, in: "Factorization and integrable systems", Oper. Theory Adv. Appl. **141**, Birkhäuser, Basel 2003, 1–102.

[22] GOHBERG I., LANCASTER P. AND RODMAN L., *Matrix polynomials*, Computer Science and Applied Mathematics, Academic Press Inc, New York, 1982.

[23] GOLUB G.H. AND VAN LOAN C.F., *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD 1996.

[24] GRASSMANN W.K., TAKSAR M.I. AND HEYMAN D.P., *Regenerative analysis and steady state distributions for Markov chains*, Oper. Res. **33** (1985), 1107–1116.

[25] GUO C.-H. *On the numerical solution of a nonlinear matrix equation in Markov chains*. Linear Algebra Appl. **288** (1999), 175–186.

[26] GUO C.-H. *Convergence analysis of the Latouche–Ramaswami algorithm for null recurrent quasi-birth-death processes*, SIAM J. Matrix Anal. Appl. **23** (2001/2), 744–760.

[27] GUO C.-H., *Comments on a shifted cyclic reduction algorithm for quasi-birth-death problems*, SIAM J. Matrix Anal. Appl. **24** (2003), 1161–1166.

[28] HE C., MEINI B. AND RHEE N.H. *A shifted cyclic reduction algorithm for quasi-birth-death problems*. SIAM J. Matrix Anal. Appl. **23** (2001/2), 673–691.

[29] LANGVILLE A.N. AND MEYER C.D., *A survey of eigenvector methods for Web information retrieval*, SIAM Rev. **47** (2005), 135–161.

[30] LI S.-Q. AND SHENG H.-D., *Generalized folding-algorithm for sojourn time analysis of finite QBD processes and its queueing applications*, Comm. Statist. Stochastic Models **12** (1996), 507–522.

[31] LATOUCHE G., *Newton's iteration for non-linear equations in Markov chains*. IMA J. Numer. Anal., **14** (1994) 583–598.

[32] LATOUCHE G. AND RAMASWAMI V., *A logarithmic reduction algorithm for quasi-birth-death processes*, J. Appl. Probab. **30** (1993), 650–674.

[33] LATOUCHE G. AND RAMASWAMI V., *Introduction to matrix analytic methods in stochastic modeling*, ASA-SIAM Series on Statistics and Applied Probability, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA 1999.

[34] MEINI B., *An improved FFT-based version of Ramaswami's formula*, Stochastic Models **13** (1997), 223–238.

[35] MEINI B., *New convergence results on functional iteration techniques for the numerical solution of $M/G/1$ type Markov chains*, Numer. Math. **78** (1997), 39–58.

[36] MOLER C., *Numerical computing with MATLAB*. `http://www.mathworks.com/moler`, Electronic edition, Section 2.11, (2004).

[37] NEUTS M.F., *Matrix-geometric solutions in stochastic models. An algorithmic approach*, Johns Hopkins Series in the Mathematical Sciences **2**, Johns Hopkins University Press, Baltimore, MD 1981.

[38] NEUTS M.F., *Structured stochastic matrices of $M/G/1$ type and their applications*, Probability: Pure and Applied **5**, Marcel Dekker Inc., New York 1989.

[39] OSTROWSKI A., *Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent. Chapitres III et IV*, Acta Math. **72** (1940), 157–257.

[40] PAGE L., BRIN S., MOTWANI R. AND WINOGRAD T., *The pagerank citation ranking: Bringing order to the Web*, `http://dbpubs.stanford.edu:8090/pub/1999-66`, (1999).

[41] RAMASWAMI V., *Matrix analytic methods for stochastic fluid flows*, in: "Proceedings of the 16th International Teletraffic Congress", Edimburgh 1999, 1019–1030.

[42] RAMASWAMI V., *Nonlinear matrix equations in applied probability—solution techniques and open problems*, SIAM Rev. **30** (1988), 256–263.

[43] RAMASWAMI V., *The generality of quasi birth-and-death processes*, in: "Advances in matrix analytic methods for stochastic models", (A. S. Alfa and S. R. Chakravarthy Eds.), Notable Publications, NJ 1998, 93–113..

[44] SERRA CAPIZZANO S., *Jordan canonical form of the Google matrix: a potential contribution to the Page Rank computation*, SIAM J. Matrix Anal. Appl. **27** 2 (2005), 305–312.

[45] STEWART G.W., *On the solution of block Hessenberg systems*, Numer. Linear Algebra Appl. **2** (1955), 287–296.

[46] STEWART W.J., *Introduction to the numerical solution of Markov chains*. Princeton University Press, Princeton, NJ 1994.

[47] VARGA R.S., *Matrix iterative analysis*, Springer Series in Computational Mathematics **27**, expanded edition, Springer-Verlag, Berlin 2000.

AMS Subject Classification: 15A24, 60J22, 60K25, 65F30

Dario A. BINI, Dipartimento di Matematica, Università di Pisa, L.go Pontecorvo 5, I-56127 Pisa, ITALY
`bini@dm.unipi.it`