New convergence results on functional iteration techniques for the numerical solution of M/G/1 type Markov chains

Beatrice Meini*
Università di Pisa

January 21, 1997

Abstract

By performing an accurate analysis of the convergence, we give a complete theoretical explanation of the experimental behaviour of functional iteration techniques for the computation of the minimal nonnegative solution G of the matrix equation $X = \sum_{i=0}^{+\infty} X^i A_i$, arising in the numerical solution of M/G/1 type Markov chains (here the A_i's are nonnegative $k \times k$ matrices such that the matrix $\sum_{i=0}^{+\infty} A_i$ is column stochastic). Moreover, we introduce a general class of functional iteration methods, which includes the standard methods, and we give an optimality convergence result in this class.

AMS subject classification. 15A51, 15A24, 60J10, 60K25, 65U05.

1 Introduction.

Let P be the infinite column stochastic matrix

$$P = \begin{pmatrix} B_1 & A_0 & & \bigcirc & \\ B_2 & A_1 & A_0 & & \\ B_3 & A_2 & A_1 & A_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \tag{1}$$

where the $k \times k$ blocks B_{i+1}, A_i, $i \geq 0$, are nonnegative and $\sum_{i=1}^{+\infty} B_i$ and $\sum_{i=0}^{+\infty} A_i$ are column stochastic matrices. Matrices of the structure (1) are known in literature as stochastic matrices of M/G/1 type [9] and arise in a wide variety of queueing problems modelled by Markov chains $\mathcal{M}$, where P^T is the transition

*Dip. Matematica, v. Buonarroti 2, 56127 Pisa, Italy. E-mail: meini@dm.unipi.it

matrix associated with $\mathcal{M}$. The computation of the probability invariant vector associated with a stochastic matrix of M/G/1 type (1), i.e., the nonnegative infinite vector π such that $\pi = P\pi$, $\|\pi\|_1 = 1$, can be reduced (compare [9]) to the computation of the minimal nonnegative solution G of the nonlinear matrix equation

$$X = \sum_{i=0}^{+\infty} X^i A_i, \tag{2}$$

where X is a $k \times k$ matrix. In the case where the matrix P is irreducible and positive recurrent [4], equation (2) has an unique nonnegative solution, which is column stochastic [9].

The problem of the computation of the nonnegative solution G of (2) has been investigated by several authors. New numerical methods have been recently proposed in [3] and in [7]: in [3] the authors propose a method based on cyclic reduction, which is numerically stable, quadratically convergent and with low computational cost; in [7] the evaluation of the matrix G is reduced to solving, by means of a doubling technique, a block Hessenberg system. However, the most commonly used algorithms for the computation of the matrix G are functional iteration methods based on the recursion

$$X_n = F(X_{n-1}), \quad n \geq 1 \tag{3}$$

where X_0 is a nonnegative matrix and $F(\cdot)$ is given by

$$F(X) = \sum_{i=0}^{+\infty} X^i A_i \tag{4}$$

or

$$F(X) = \left(A_0 + \sum_{i=2}^{+\infty} X^i A_i \right) (I - A_1)^{-1} \tag{5}$$

or

$$F(X) = A_0 \left(I - \sum_{i=1}^{+\infty} X^{i-1} A_i \right)^{-1}. \tag{6}$$

The sequence $\{X_n\}$ of (3), with $F(\cdot)$ defined by (4), (5) or (6), converges linearly to the matrix G. However a detailed analysis of the speed of convergence, in relation with the particular function $F(\cdot)$ or with the starting matrix X_0, has not yet been provided. In [5] Latouche shows that, in the case where $X_0 = O$, the sequence (3), where $F(\cdot)$ is (4), or (5), or (6), converges monotonically to the matrix G; moreover, he proves that the method based on (6) is faster than that based on (5), and the method based on (5) is faster than that based on (4), but the rate of convergence of the three sequences is not estimated. In [5] the author also shows that, by choosing as starting matrix X_0 a general column stochastic matrix, the sequence (3), where $F(\cdot)$ is (4), or (5), or (6), is a sequence of stochastic matrices that still converges to the matrix G. Moreover, Latouche observes [6] that, in many numerical experimentations, the sequence obtained in this way converges much faster than the sequence obtained by starting with $X_0 = O$.

In this paper we introduce a class of functional iterations (3) containing the ones described above; then we associate, with each function F defining a functional iteration of this class, a suitable nonnegative $k \times k$ matrix R_F. In the case where the starting matrix of (3) is $X_0 = O$, we first show that the sequence (3) converges monotonically to the matrix G. Then we prove that, for the monotonic convergence of the matrices $\{X_n\}$ to the solution G, the spectral radius ρ_F of the matrix R_F coincides with the mean asymptotic rate of convergence of the sequence $\{X_n\}_n$. On the other hand, for each function $F(\cdot)$ belonging to the above class, the matrix R_F can be obtained as a regular splitting (compare [10]) of the same M-matrix. This fact allows us to prove that the mean asymptotic rates of convergence r_N, r_T, r_U of the methods based on (4), (5), (6), respectively, are related by the following inequality:

$$r_U \leq r_T \leq r_N.$$

Moreover, we prove that the method based on (6) is the fastest method in the class of functional iterations that we introduced.

In the case where the starting matrix X_0 of (3) is any column stochastic matrix, we associate with any function $F(\cdot)$ of the above class a nonnegative $k^2 \times k^2$ matrix $\widehat{R}_F$. The increasing of the speed of convergence of the sequence $\{X_n\}$, with respect to the case $X_0 = O$ observed in [6], can be formally explained by means of the matrices $\widehat{R}_F$. In fact, we first observe that the spectral radius of the matrix $\widehat{R}_F$ coincides with the spectral radius of the matrix R_F. Then we show that, at each step n, the vector $\hat{\mathbf{e}}_n$ obtained by the column-wise arrangement of the entries of the matrix $G - X_n$, which represents the error at step n, is orthogonal to the nonnegative left eigenvector $\mathbf{w}^T$ associated with the spectral radius of the matrix $\widehat{R}_F$. This fact allows us to prove that the asymptotic rate of convergence r of the sequence $\{X_n\}$ coincides with the spectral radius of the matrix W_F obtained by projecting the matrix $\widehat{R}_F$ onto a suitable linear space E. The linear space E is such that $\mathbf{w}$ is orthogonal to any vector belonging to E. Whence, in the case where the matrix $\widehat{R}_F$ is irreducible the spectral radius of the matrix W_F is less than or equal to the second largest modulus eigenvalue of the matrix $\widehat{R}_F$; in the case where the matrix $\widehat{R}_F$ is reducible, the spectral radius of the matrix W_F is less than or equal to the spectral radius of $\widehat{R}_F$.

The paper is organized as follows. In section 2 we analyze functional iteration methods (3) defined by (4), (5) and (6) with $X_0 = O$. In section 3 we introduce the general class of functional iteration methods and we estimate the rate of convergence of the sequences generated in this class, in the case where $X_0 = O$. In section 4 we analyze functional iteration methods having, as starting point X_0, a column stochastic matrix.

In what follows, $\mathbf{e}^T$ denotes the k-dimensional vector $\mathbf{e}^T = (1, 1, \ldots, 1)$, $\rho(M)$ denotes the spectral radius of the matrix M and $\bar{A}_i$, for $i \geq 1$, denotes the $k \times k$ matrix $\bar{A}_i = \sum_{j=i}^{+\infty} G^{j-i} A_i$. We also introduce the vector norm $\|\mathbf{v}\|_1 = \sum_i |v_i|$, for vectors $\mathbf{v} = \{v_i\}_i$ and the corresponding induced matrix norm $\|A\|_1 = \max_j \sum_i |a_{i,j}|$, for matrices $A = \{a_{i,j}\}_{i,j}$. We suppose that the matrix P is irreducible and positive recurrent, so that the matrix equation (2) has a unique nonnegative solution. Moreover we suppose that the sequence $\{X_n\}$ of (3) does not reach the solution G in a finite number of steps.

2 Convergence properties of standard functional iteration methods.

In this section we analyze functional iteration methods defined by (3), with $X_0 = O$ and $F(X)$ given by (4), (5), (6). A convergence analysis of such iterative methods have already been performed by G. Latouche in [5]. More precisely, by following the notations introduced in [5], let us denote by $\{X_n^{(N)}\}$ the sequence defined by (4) (Natural algorithm), by $\{X_n^{(T)}\}$ the sequence defined by (5) (Traditional algorithm) and by $\{X_n^{(U)}\}$ the sequence defined by (6) (algorithm based on the matrix $U = \sum_{i=1}^{+\infty} G^{i-1} A_i$), with $X_0^{(N)} = X_0^{(T)} = X_0^{(U)} = O$. Then the following result holds [5]:

Theorem 1 *The sequences $\{X_n^{(N)}\}_n$, $\{X_n^{(T)}\}_n$, $\{X_n^{(U)}\}_n$ converge monotonically to the matrix G. Moreover, for any $n \geq 0$, it holds*

$$X_n^{(N)} \leq X_n^{(T)} \leq X_n^{(U)}.$$

Theorem 1 allows us to deduce [5] that the number of iterations I_N, I_T, I_U sufficient to obtain the same approximation of the matrix G by means of the sequences $\{X_n^{(N)}\}_n$, $\{X_n^{(T)}\}_n$, $\{X_n^{(U)}\}_n$, respectively, are such that

$$I_U \leq I_T \leq I_N.$$

Therefore the method based on (6) is faster than that based on (5), and the method based on (5) is faster than that based on (4). We give a more precise result, by estimating also the rate of convergence of the three sequences. Let us first prove the following

Lemma 2 *Let X, Y be $k \times k$ matrices. Then, for any integer $i \geq 1$, it holds that*

$$X^i - Y^i = \sum_{j=1}^{i} X^{j-1}(X - Y)Y^{i-j}.$$

Proof. The equality can be simply proved by induction on i. $\square$

Define, for each integer n, the matrices $E_n^{(N)} = G - X_n^{(N)}$, $E_n^{(T)} = G - X_n^{(T)}$ and $E_n^{(U)} = G - X_n^{(U)}$, which represent the error at step n for the sequences (3) defined by (4), (5), (6), respectively, with $X_0 = O$. Observe that, for the monotonicity stated by theorem 1, the errors at any step n are such that $E_n^{(N)} \geq O$, $E_n^{(T)} \geq O$, $E_n^{(U)} \geq O$. Moreover it is possible to express them recursively, as shown by the following

Proposition 3 *For every integer $n \geq 0$ it holds that*

$$\mathbf{e}^T E_{n+1}^{(N)} = \mathbf{e}^T E_n^{(N)} R_n^{(N)}, \quad \mathbf{e}^T E_{n+1}^{(T)} = \mathbf{e}^T E_n^{(T)} R_n^{(T)}, \quad \mathbf{e}^T E_{n+1}^{(U)} = \mathbf{e}^T E_n^{(U)} R_n^{(U)}$$

where, for $n \geq 0$,

$$\begin{aligned} R_n^{(N)} &= \textstyle\sum_{i=1}^{+\infty} \sum_{j=1}^{i} {X_n^{(N)}}^{i-j} A_i \\ R_n^{(T)} &= \left(\textstyle\sum_{i=1}^{+\infty} \sum_{j=1}^{i} {X_n^{(T)}}^{i-j} A_i - A_1 \right) (I - A_1)^{-1} \\ R_n^{(U)} &= \left(\textstyle\sum_{i=2}^{+\infty} \sum_{j=2}^{i} {X_n^{(U)}}^{i-j} A_i \right) \left(I - \textstyle\sum_{i=1}^{+\infty} {X_n^{(U)}}^{i-1} A_i \right)^{-1}. \end{aligned} \tag{7}$$

Proof. Let us first analyze the functional iteration method based on (4). From lemma 2 and from (2) and (3) it follows that

$$E_{n+1}^{(N)} = G - X_{n+1}^{(N)} = \sum_{i=1}^{+\infty} \left(G^i - {X_n^{(N)}}^i \right) A_i = \sum_{i=1}^{+\infty} \sum_{j=1}^{i} G^{j-1} E_n^{(N)} {X_n^{(N)}}^{i-j} A_i.$$

Whence, since G is column stochastic, we have

$$\mathbf{e}^T E_{n+1}^{(N)} = \mathbf{e}^T \sum_{i=1}^{+\infty} \sum_{j=1}^{i} E_n^{(N)} {X_n^{(N)}}^{i-j} A_i = \mathbf{e}^T E_n^{(N)} R_n^{(N)}.$$

Analogously, for the functional iteration method based on (5), we obtain

$$\begin{aligned} E_{n+1}^{(T)} = G - X_{n+1}^{(T)} = \textstyle\sum_{i=2}^{+\infty} \left(G^i - {X_n^{(T)}}^i \right) A_i (I - A_1)^{-1} = \\ \textstyle\sum_{i=2}^{+\infty} \sum_{j=1}^{i} G^{j-1} E_n^{(T)} {X_n^{(T)}}^{i-j} A_i (I - A_1)^{-1}. \end{aligned}$$

Hence, from the stochasticity of G, we obtain

$$\mathbf{e}^T E_{n+1}^{(T)} = \mathbf{e}^T E_n^{(T)} \left(\sum_{i=2}^{+\infty} \sum_{j=1}^{i} {X_n^{(T)}}^{i-j} A_i \right) (I - A_1)^{-1} = \mathbf{e}^T E_n^{(T)} R_n^{(T)}.$$

Let us now analyze the functional iteration method based on (6). From (3) and (2) we have

$$
\begin{aligned}
&E_{n+1}^{(U)} = \\
&G - X_{n+1}^{(U)} = A_0 \left[\left(I - \textstyle\sum_{i=1}^{+\infty} G^{i-1} A_i \right)^{-1} - \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i \right)^{-1} \right] = \\
&G \left[I - \left(I - \textstyle\sum_{i=1}^{+\infty} G^{i-1} A_i \right) \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i \right)^{-1} \right].
\end{aligned}
$$

Whence, from lemma 2, it follows that

$$
\begin{aligned}
&E_{n+1}^{(U)} = \\
&G \left[I - \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i - \sum_{i=2}^{+\infty} \sum_{j=1}^{i-1} G^{j-1} E_n^{(U)} X_n^{(U)^{i-1-j}} A_i \right) \cdot \right. \\
&\qquad\qquad \left. \cdot \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i \right)^{-1} \right] = \\
&G \left(\textstyle\sum_{i=2}^{+\infty} \sum_{j=1}^{i-1} G^{j-1} E_n^{(U)} X_n^{(U)^{i-1-j}} A_i \right) \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i \right)^{-1}.
\end{aligned}
$$

Therefore, from the stochasticity of G, we have

$$
\begin{aligned}
&\mathbf{e}^T E_{n+1}^{(U)} = \\
&\mathbf{e}^T E_n^{(U)} \left(\textstyle\sum_{i=2}^{+\infty} \sum_{j=1}^{i-1} X_n^{(U)^{i-1-j}} A_i \right) \left(I - \textstyle\sum_{i=1}^{+\infty} X_n^{(U)^{i-1}} A_i \right)^{-1} = \mathbf{e}^T E_n^{(U)} R_n^{(U)}
\end{aligned}
$$

□

Proposition 3 allows us to express the norm of the error at step n, by means of the matrices $R_n^{(N)}$, $R_n^{(T)}$, $R_n^{(U)}$ of (7):

Corollary 4 *For every integer $n \geq 1$ it holds that*

$$
\left\| E_n^{(N)} \right\|_1 = \left\| \prod_{i=0}^{n-1} R_i^{(N)} \right\|_1, \quad \left\| E_n^{(T)} \right\|_1 = \left\| \prod_{i=0}^{n-1} R_i^{(T)} \right\|_1, \quad \left\| E_n^{(U)} \right\|_1 = \left\| \prod_{i=0}^{n-1} R_i^{(U)} \right\|_1,
$$

where the matrices $R_i^{(N)}$, $R_i^{(T)}$, $R_i^{(U)}$ are defined in (7) and $\prod_{i=0}^{n-1} H_i = H_0 H_1 \ldots H_{n-1}$, for $k \times k$ matrices H_i.

Proof. For every integer n, since $E_n^{(N)} \geq O$, we have $\|E_n^{(N)}\|_1 = \|\mathbf{e}^T E_n^{(N)}\|_1$. Whence, for proposition 3, $\|E_n^{(N)}\|_1 = \|\mathbf{e}^T E_0^{(N)} \prod_{i=0}^{n-1} R_i^{(N)}\|_1$. On the other hand, since $X_0^{(N)} = O$, we have $\mathbf{e}^T E_0^{(N)} = \mathbf{e}^T G = \mathbf{e}^T$. Therefore, $\|E_n^{(N)}\|_1 = \|\mathbf{e}^T \prod_{i=0}^{n-1} R_i^{(N)}\|_1 = \| \prod_{i=0}^{n-1} R_i^{(N)}\|_1$. The analogous relations for $E_n^{(T)}$ and $E_n^{(U)}$ readily follow by using the same arguments. □

It is also possible to estimate the asymptotic rate of convergence of the sequences $\{X_n^{(N)}\}$, $\{X_n^{(T)}\}$, $\{X_n^{(U)}\}$. Let us first define the matrices

$$\begin{aligned} R^{(N)} &= \textstyle\sum_{i=1}^{+\infty} \bar{A}_i \\ R^{(T)} &= \left(\textstyle\sum_{i=1}^{+\infty} \bar{A}_i - A_1\right)(I - A_1)^{-1} \\ R^{(U)} &= \left(\textstyle\sum_{i=2}^{+\infty} \bar{A}_i\right)\left(I - \bar{A}_1\right)^{-1} \end{aligned} \tag{8}$$

such that

$$R^{(N)} = \lim_n R_n^{(N)}, \; R^{(T)} = \lim_n R_n^{(T)}, \; R^{(U)} = \lim_n R_n^{(U)}. \tag{9}$$

Then the following theorem holds:

Theorem 5 *Let*

$$r_N = \lim_n \sqrt[n]{\|E_n^{(N)}\|}, \quad r_T = \lim_n \sqrt[n]{\|E_n^{(T)}\|}, \quad r_U = \lim_n \sqrt[n]{\|E_n^{(U)}\|},$$

be the mean asymptotic rates of convergence for the sequences $\{X_n^{(N)}\}$, $\{X_n^{(T)}\}$, $\{X_n^{(U)}\}$, *respectively, where* $\|\cdot\|$ *is any matrix norm. Then it holds that*

$$r_N = \rho(R^{(N)}), \; r_T = \rho(R^{(T)}), \; r_U = \rho(R^{(U)}),$$

where the matrices $R^{(N)}$, $R^{(T)}$, $R^{(U)}$ *are defined in (8).*

Proof. We prove the theorem only for the sequence $\{X_n^{(N)}\}$ and we leave the reader to complete the proof for the sequences $\{X_n^{(T)}\}$ and $\{X_n^{(U)}\}$. For corollary 4 it holds that

$$r_N = \lim_n \sqrt[n]{\|E_n^{(N)}\|_1} = \lim_n \sqrt[n]{\left\| \prod_{i=0}^{n-1} R_i^{(N)} \right\|_1}. \tag{10}$$

For the monotonicity of the sequence $\{X_n^{(N)}\}$ stated by theorem 1 and for the properties of the spectral radius (compare [10]), we find

$$\lim_n \sqrt[n]{\left\| \prod_{i=0}^{n-1} R_i^{(N)} \right\|_1} \leq \lim_n \sqrt[n]{\|{R^{(N)}}^n\|_1} = \rho(R^{(N)}),$$

whence

$$r_N \leq \rho(R^{(N)}). \tag{11}$$

Let us now prove the opposite inequality. Suppose first that the nonnegative vector $\mathbf{e}^T R_0^{(N)}$ has no null components. For any integer $m < n$, for the monotonicity of the sequence $\{X_n^{(N)}\}$ and for the monotonicity of $\|\cdot\|_1$, it holds that

$$\left\| \prod_{i=0}^{n-1} R_i^{(N)} \right\|_1 \geq \left\| \prod_{i=0}^{m-1} R_i^{(N)} {R_m^{(N)}}^{n-m} \right\|_1 \geq \left\| {R_0^{(N)}}^m {R_m^{(N)}}^{n-m} \right\|_1. \tag{12}$$

Hence, for any integer m, since the nonnegative vector $\mathbf{e}^T {R_0^{(N)}}^m$ also has no null components, it holds that

$$\left\| {R_0^{(N)}}^m {R_m^{(N)}}^{n-m} \right\|_1 \geq c_m \left\| {R_m^{(N)}}^{n-m} \right\|_1 \tag{13}$$

where $c_m > 0$ is a suitable constant. From (9) it follows that, for any fixed $\epsilon > 0$, there exists an integer m_0 such that $\rho(R_{m_0}^{(N)}) \geq \rho(R^{(N)}) - \epsilon$. Whence, from (10), (12) and (13), we obtain

$$\begin{aligned} r_N \geq \lim_n \sqrt[n]{\left\| {R_0^{(N)}}^{m_0} {R_{m_0}^{(N)}}^{n-m_0} \right\|_1} \geq \lim_n \sqrt[n]{c_{m_0} \left\| {R_{m_0}^{(N)}}^{n-m_0} \right\|_1} \geq \\ \lim_n \sqrt[n]{c_{m_0}} \sqrt[n]{\left\| {R_{m_0}^{(N)}}^{n} \right\|_1} \sqrt[n]{\left\| {R_{m_0}^{(N)}}^{m_0} \right\|_1^{-1}} = \lim_n \sqrt[n]{\left\| {R_{m_0}^{(N)}}^{n} \right\|_1} = \\ \rho(R_{m_0}^{(N)}) \geq \rho(R^{(N)}) - \epsilon. \end{aligned}$$

Therefore, for the arbitrarity of ϵ, the above inequality, together with (11), leads to the thesis. Suppose now that the first i components of the nonnegative vector $\mathbf{e}^T R_0^{(N)} = \mathbf{e}^T E_1^{(N)}$ are zero and that the remaining components are nonnull. Whence, for the monotonic convergence of the sequence $\{X_n^{(N)}\}$, it follows that also the first i components of the nonnegative vector $\mathbf{e}^T E_n^{(N)}$ are zero for any $n \geq 1$. Moreover, it can easily be proved (we leave the proof to the reader) that, for any $n \geq 0$, the first i components of the vector $\mathbf{e}^T R_n^{(N)}$ are zero. Therefore, the matrices $R_n^{(N)}$ have the structure

$$R_n^{(N)} = \begin{pmatrix} O & T_n^{(N)} \\ O & U_n^{(N)} \end{pmatrix}$$

where $U_n^{(N)}$ is a $(k-i) \times (k-i)$ matrix and $T_n^{(N)}$ is an $i \times (k-i)$ matrix. Moreover, the matrix $R^{(N)}$ has the structure

$$R^{(N)} = \begin{pmatrix} O & T^{(N)} \\ O & U^{(N)} \end{pmatrix}$$

where $U^{(N)} = \lim_n U_n^{(N)}$, $T^{(N)} = \lim_n T_n^{(N)}$. If, except for the first i components, no other entry of the vector $\mathbf{e}^T E_n^{(N)}$ vanishes in a finite number of steps, then the matrices $U_n^{(N)}$ have no null columns for any $n \geq 1$. Let $\mathbf{v}_1^T$ be the $k - i$

dimensional vector defined by the nonnull components of the vector $\mathbf{e}^T E_1^{(N)}$. From the monotonicity of the sequence $\{X_n^{(N)}\}$ and from the monotonicity of $\|\cdot\|_1$, it follows that, for any integer $m < n$

$$\begin{array}{l} \left\| \prod_{i=0}^{n-1} R_i^{(N)} \right\|_1 = \left\| \mathbf{v}_1^T \prod_{i=1}^{n-1} U_i^{(N)} \right\|_1 \geq c \left\| \prod_{i=1}^{n-1} U_i^{(N)} \right\|_1 \geq \\ \quad c \left\| \prod_{i=1}^{m-1} U_i^{(N)} {U_m^{(N)}}^{n-m} \right\|_1 \geq c \left\| {U_1^{(N)}}^{m-1} {U_m^{(N)}}^{n-m} \right\|_1 \geq c_m \left\| {U_m^{(N)}}^{n-m} \right\|_1, \end{array}$$

where $c, c_m > 0$ are suitable constants. Whence, for every $\epsilon > 0$, there exists m_0 such that

$$r_N \geq \lim_n \sqrt[n]{c_{m_0} \left\| {U_{m_0}^{(N)}}^{n-m_0} \right\|_1} = \rho(U_{m_0}^{(N)}) \geq \rho(U^{(N)}) - \epsilon = \rho(R^{(N)}) - \epsilon,$$

which leads to the thesis, for the arbitrarity of ϵ. If the components $i+1, \ldots, i+i_0$ of the vector $\mathbf{e}^T E_n^{(N)}$ vanish in a finite number of steps and no other component vanishes, it follows that, for $n \geq 1$, the matrices $U_n^{(N)}$ have the structure

$$U_n^{(N)} = \begin{pmatrix} \widehat{V}_n^{(N)} & \widehat{T}_n^{(N)} \\ O & \widehat{U}_n^{(N)} \end{pmatrix}$$

where $\widehat{U}_n^{(N)}$ is a $(k-i-i_0) \times (k-i-i_0)$ matrix having no null columns and $\widehat{V}_n^{(N)}$ is an upper triangular matrix with null diagonal entries. Therefore we apply the same arguments to the matrix $\widehat{U}_n^{(N)}$, concluding the proof. $\square$

The following proposition, together with theorem 5, allows us to compare the rate of convergence of the three sequences. In particular, it shows that the sequence $\{X_n^{(U)}\}$ based on (6) is asymptotically faster than the sequences $\{X_n^{(N)}\}$ and $\{X_n^{(T)}\}$ based on (4) and (5), respectively:

Proposition 6 *The spectral radii of the matrices $R^{(N)}$, $R^{(T)}$ and $R^{(U)}$ are related by the following inequality*

$$\rho(R^{(U)}) \leq \rho(R^{(T)}) \leq \rho(R^{(N)}).$$

Proof. Observe that

$$R^{(N)} = C^{(N)} {D^{(N)}}^{-1}, \; R^{(T)} = C^{(T)} {D^{(T)}}^{-1}, \; R^{(U)} = C^{(U)} {D^{(U)}}^{-1},$$

where

$$\begin{array}{ll} C^{(N)} = \sum_{i=1}^{+\infty} \bar{A}_i, & D^{(N)} = I \\ C^{(T)} = \sum_{i=1}^{+\infty} \bar{A}_i - A_1, & D^{(T)} = I - A_1 \\ C^{(U)} = \sum_{i=2}^{+\infty} \bar{A}_i, & D^{(U)} = I - \bar{A}_1 \end{array}$$

are such that

$$D^{(N)} - C^{(N)} = D^{(T)} - C^{(T)} = D^{(U)} - C^{(U)} = I - \sum_{i=1}^{+\infty} \bar{A}_i,$$

that is, the matrices $R^{(N)}$, $R^{(T)}$ and $R^{(U)}$ are obtained by means of regular splittings (compare [10]) of the M-matrix $I-\sum_{i=1}^{+\infty}\bar{A}_i$. Whence, for the Perron-Frobenius theorem (compare [2]), since

$$C^{(U)} \leq C^{(T)} \leq C^{(N)},$$

it follows that

$$\rho(R^{(U)}) \leq \rho(R^{(T)}) \leq \rho(R^{(N)}).$$

□

3 A general class of functional iteration methods.

Functional iteration methods (3) defined by (4), (5), (6) belong to a more general class where the function $F(\cdot)$ is such that the matrix G satisfies the equation $X = F(X)$. More precisely, consider the formal power series $A(z) = \sum_{i=0}^{+\infty} z^i A_i$ (observe that $A(z)$ converges in the closed unit circle) and suppose that $A(z) = zH(z) + K(z)$, where $H(z) = \sum_{i=0}^{+\infty} z^i H_i$ and $K(z) = \sum_{i=0}^{+\infty} z^i K_i$ are such that $H_i \geq O$, $K_i \geq O$, $i \geq 0$ (observe that $A_0 = K_0$ and $A_i = H_{i-1} + K_i$, for $i \geq 1$). Then the following property holds

Proposition 7 *The matrix G solves equation (2) if and only if it solves the matrix equation*

$$X = K(X)(I - H(X))^{-1} \tag{14}$$

where X is a $k \times k$ matrix and $K(X) = \sum_{i=0}^{+\infty} X^i K_i$, $H(X) = \sum_{i=0}^{+\infty} X^i H_i$.

Proof. The matrix G solves equation (2) if and only if it solves equation

$$X(I - H(X)) = K(X).$$

Since $H_i \leq A_{i+1}$ for $i \geq 0$ and since the matrix $I - \sum_{i=0}^{+\infty} G^i A_{i+1}$ is nonsingular (compare [9]), it follows that $I - H(G)$ is nonsingular. Hence G solves (14). □

Equation (14) allows us to define the sequence of matrices

$$X_n = F(X_{n-1}), \quad n \geq 1$$

where

$$F(X) = K(X)(I - H(X))^{-1}. \tag{15}$$

By choosing $H(z) = 0$, $H(z) = A_1$, $H(z) = \sum_{i=0}^{+\infty} z^i A_{i+1}$, we obtain the functional iteration methods defined by (4), (5), (6), respectively. The results stated in section 2 can easily be extended to iterations (3) defined by (15), with $X_0 = O$. Let us first prove the following result, which extends theorem 1:

Theorem 8 *The sequence $X_n = F(X_{n-1})$, $X_0 = O$, where $F(\cdot)$ is given by (15), converges monotonically to the matrix G.*

Proof. Let us prove the monotonicity of the sequence $\{X_n\}$ by induction on n. For $n = 1$ the thesis is true, since $X_1 = F(X_0) \geq O = X_0$. Suppose that $X_n \geq X_{n-1}$ and let us prove that $X_{n+1} \geq X_n$: from (15) and from the inductive hypothesis, we obtain

$$X_{n+1} = F(X_n) = K(X_n)(I - H(X_n))^{-1} \geq K(X_{n-1})(I - H(X_{n-1}))^{-1} = X_n.$$

$\square$

Define, for each integer $n \geq 0$, the matrices $E_n = G - X_n$, which represent the error at step n of the sequence $\{X_n\}$. Then the following general recursive relations for the matrices $\{E_n\}$ hold:

Proposition 9 *Let $X_n = F(X_{n-1})$, $X_0 \geq O$, where $F(\cdot)$ is defined by (15). Then for every integer $n \geq 0$ it holds that*

$$E_{n+1} = \left(\sum_{i=1}^{+\infty} \sum_{j=1}^{i} (G^{j-1} E_n X_n^{i-j} K_i + G^j E_n X_n^{i-j} H_i) \right) (I - H(X_n))^{-1} \qquad (16)$$

or, in equivalent form,

$$E_{n+1} = \left(E_n \sum_{i=1}^{+\infty} X_n^{i-1} K_i + \sum_{j=2}^{+\infty} G^{j-1} E_n \sum_{i=j}^{+\infty} X_n^{i-j} A_i \right) (I - H(X_n))^{-1}. \qquad (17)$$

Proof. From (3) and (15) it follows that

$$E_{n+1} = G - X_{n+1} = K(G)(I - H(G))^{-1} - K(X_n)(I - H(X_n))^{-1} =$$

$$\textstyle\sum_{i=0}^{+\infty} G^i K_i (I - H(G))^{-1} - \sum_{i=0}^{+\infty} X_n^i K_i (I - H(X_n))^{-1} =$$

$$\textstyle\sum_{i=0}^{+\infty} (G^i - X_n^i) K_i (I - H(X_n))^{-1} + K(G) \left((I - H(G))^{-1} - (I - H(X_n))^{-1} \right) =$$

$$\textstyle\sum_{i=0}^{+\infty} (G^i - X_n^i) K_i (I - H(X_n))^{-1} + G \left(I - (I - H(G))(I - H(X_n))^{-1} \right).$$

Whence, from lemma 2

$$\textstyle\sum_{i=0}^{+\infty} (G^i - X_n^i) K_i (I - H(X_n))^{-1} =$$

$$\textstyle\sum_{i=1}^{+\infty} \sum_{j=1}^{i} G^{j-1} (G - X_n) X_n^{i-j} K_i (I - H(X_n))^{-1}$$

and

$I-(I-H(G))(I-H(X_n))^{-1}=$

$I-\left(I-\sum_{i=0}^{+\infty}X_n^iH_i-\sum_{i=1}^{+\infty}\sum_{j=1}^{i}G^{j-1}(G-X_n)X_n^{i-j}H_i\right)(I-H(X_n))^{-1}=$

$\left(\sum_{i=1}^{+\infty}\sum_{j=1}^{i}G^{j-1}(G-X_n)X_n^{i-j}H_i\right)(I-H(X_n))^{-1}.$

Therefore

$$E_{n+1}=\left(\sum_{i=1}^{+\infty}\sum_{j=1}^{i}(G^{j-1}E_nX_n^{i-j}K_i+G^jE_nX_n^{i-j}H_i)\right)(I-H(X_n))^{-1}.$$

From the above equation and from the relation $A_i=H_{i-1}+K_i$, $i\geq 1$, we obtain formula (17). □

From the above proposition the generalization of proposition 3 follows:

Corollary 10 *Let $X_n=F(X_{n-1})$, $X_0=O$, where $F(\cdot)$ is defined by (15). Then for every integer $n\geq 0$ it holds that*

$$\mathbf{e}^TE_{n+1}=\mathbf{e}^TE_nR_n \tag{18}$$

where, for $n\geq 0$,

$$R_n=\left(\sum_{j=1}^{+\infty}\sum_{i=j}^{+\infty}X_n^{i-j}A_i-H(X_n)\right)(I-H(X_n))^{-1}. \tag{19}$$

Proof. For the stochasticity of the matrix G, from (16) and from the relation $A_i=H_{i-1}+K_i$ we obtain

$\mathbf{e}^TE_{n+1}=$

$\mathbf{e}^TE_n\left(\sum_{i=1}^{+\infty}\sum_{j=1}^{i}X_n^{i-j}(K_i+H_i)\right)(I-H(X_n))^{-1}=$

$\mathbf{e}^TE_n\left(\sum_{j=1}^{+\infty}\sum_{i=j}^{+\infty}X_n^{i-j}A_i-H(X_n)\right)(I-H(X_n))^{-1}.$

□

The following result is a straightforward consequence of corollary 10 and a generalization of corollary 4 (we leave the proof to the reader):

Corollary 11 *For every integer $n\geq 1$ it holds that*

$$\left\|E_n\right\|_1=\left\|\prod_{i=0}^{n-1}R_i\right\|_1 \tag{20}$$

where the matrices R_n are defined in (19) and $\prod_{i=0}^{n-1}R_i=R_0R_1\dots R_{n-1}$.

Define the matrix

$$R_F = \left(\sum_{i=1}^{+\infty} \bar{A}_i - H(G) \right) (I - H(G))^{-1} \tag{21}$$

such that $\lim_n R_n = R_F$. The matrix R_F allows us to express the asymptotic rate of convergence of the sequence $\{X_n\}$, by means of the following

Theorem 12 *Let*

$$r = \lim_n \sqrt[n]{\|E_n\|}$$

be the mean asymptotic rate of convergence of the sequence $X_n = F(X_{n-1})$, $X_0 = O$, *where* $F(\cdot)$ *is defined in (15) and* $\|\cdot\|$ *is any matrix norm. Then*

$$r = \rho(R_F)$$

where R_F *is given in (21).*

Proof. The theorem can easily be proved by using the same arguments of theorem 5. □

From the above theorem it follows that the speed of convergence of the sequence $X_n = F(X_{n-1})$, $X_0 = O$, defined by (15) is related to the spectral radius of the matrix R_F of (21). Observe that the matrix R_F can be written in the form

$$R_F = CD^{-1},$$

where

$$C = \sum_{i=1}^{+\infty} \bar{A}_i - H(G), \quad D = I - H(G)$$

are such that

$$C \geq O, \quad D^{-1} \geq O, \quad D - C = I - \sum_{i=1}^{+\infty} \bar{A}_i,$$

that is, R_F is obtained by means of a regular splitting (compare [10]) of the M-matrix $I - \sum_{i=1}^{+\infty} \bar{A}_i$. Suppose that $R_1 = C_1 D_1^{-1}$ and $R_2 = C_2 D_2^{-1}$ are the matrices of (21) associated with the functions $F_1(\cdot)$, $F_2(\cdot)$ of (15), respectively, and suppose that $C_1 \leq C_2$. From the Perron-Frobenius theorem (compare [2]), it follows that

$$\rho(R_1) \leq \rho(R_2).$$

On the other hand observe that, for any choice of $F(\cdot)$ defined by (15), since $H_i \geq 0$ for any i, it holds that

$$H(G) \leq \sum_{i=1}^{+\infty} G^{i-1} A_i = \bar{A}_1.$$

Whence

$$C = \sum_{i=1}^{+\infty} \bar{A}_i - H(G) \geq \sum_{i=2}^{+\infty} \bar{A}_i = C^{(U)},$$

where $R^{(U)} = C^{(U)} {D^{(U)}}^{-1}$ is the matrix associated with the sequence $\{X_n^{(U)}\}$ defined by (6). Therefore, the functional iteration method defined by (6) is the fastest one in the class of method defined by (15), with $X_0 = O$.

4 Functional iteration methods generating a sequence of stochastic matrices.

Consider functional iteration methods

$$X_n = F(X_{n-1}), \quad X_0 = Q \tag{22}$$

where $F(\cdot)$ is given by (15) and Q is a column stochastic matrix. The sequence $\{X_n\}$ is a sequence of stochastic matrices that still converges to the matrix G (compare [5]). We prove that such sequence of matrices converges faster than the sequence of matrices obtained by starting with $X_0 = O$.

Proposition 9 expresses recursively the error E_n of the sequence $X_n = F(X_{n-1})$, where $F(\cdot)$ is defined in (15) and X_0 is any nonnegative matrix. In the case where $X_0 = O$, since at each step n it holds $X_n \leq G$, we easily obtain formula (20) by left multiplying (17) by $\mathbf{e}^T$. In the case where X_0 is not necessarily equal to zero (hence E_n is not necessarily a nonnegative matrix), in order to express the error at step n in a simpler form, we introduce the matrix tensor product $\otimes$ such that $A \otimes B$ is the block matrix defined by the blocks $a_{i,j}B$, where $A = \{a_{i,j}\}_{i,j}$. In fact, if we represent a matrix A by the vector $\mathbf{a}$ obtained by column-wise arranging the entries of A, we find that the relation AXB can be rewritten as $(B^T \otimes A)\mathbf{x}$, where A, X, B are matrices of compatible sizes. Hence, the recursive formula (17) can be rewritten as

$$\hat{\mathbf{e}}_{n+1} = \widehat{R}_n \hat{\mathbf{e}}_n \tag{23}$$

where $\hat{\mathbf{e}}_n$, for $n \geq 0$, is the k^2-dimensional vector associated with the matrix E_n and

$$\begin{aligned}\widehat{R}_n = \textstyle\sum_{i=1}^{+\infty} \left(X_n^{i-1} K_i (I - H(X_n))^{-1}\right)^T \otimes I + \\ \textstyle\sum_{j=2}^{+\infty} \sum_{i=j}^{+\infty} \left(X_n^{i-j} A_i (I - H(X_n))^{-1}\right)^T \otimes G^{j-1},\end{aligned}$$

where I denotes the $k \times k$ identity matrix. Observe that the matrix $\widehat{R}_n$, for $n \geq 0$, can be expressed in the form

$$\widehat{R}_n = \sum_{j=0}^{+\infty} Y_{j,n} \otimes G^j \tag{24}$$

where

$$\begin{array}{l} Y_{0,n} = \sum_{i=1}^{+\infty} \left(X_n^{i-1} K_i (I - H(X_n))^{-1}\right)^T \\ Y_{j,n} = \sum_{i=j}^{+\infty} \left(X_n^{i-j} A_{i+1} (I - H(X_n))^{-1}\right)^T, \quad j \geq 1. \end{array}$$

Let $Y_j = \lim_n Y_{j,n}$, $j \geq 0$, and let $Y = \sum_{j=0}^{+\infty} Y_j$. Observe that

$$Y = \left(\Big(\sum_{j=1}^{+\infty} \sum_{i=j}^{+\infty} G^{i-j} A_i - H(G) \Big) (I - H(G))^{-1} \right)^T = R_F^T,$$

where $R_F = \lim_n R_n$ of (19) and is related to the speed of convergence of the sequence $X_n = F(X_{n-1})$, $X_0 = O$. From the above equation it follows, in particular, that $\rho(Y) = \rho(R_F)$. We prove also that

Proposition 13 *The spectral radius of the matrix $\widehat{R}_F = \lim_n \widehat{R}_n$ is given by*

$$\rho(\widehat{R}_F) = \rho(R_F).$$

Moreover, a nonnegative k^2-dimensional vector $\mathbf{w}$ such that

$$\mathbf{w}^T \widehat{R}_F = \lambda \mathbf{w}^T,$$

where $\lambda = \rho(R_F)$, is given by $\mathbf{w}^T = (\mathbf{y}^T \otimes \mathbf{e}^T)$, where $\mathbf{y}$ is a nonnegative k-dimensional vector such that $R_F \mathbf{y} = \lambda \mathbf{y}$.

Proof. Let S be the Schur canonical form [10] of the matrix G. Then, the matrix $T = \sum_{j=0}^{+\infty} (Y_j \otimes S^j)$ is similar to the matrix $\widehat{R}_F$. Moreover, due to the properties of the matrix S, the eigenvalues of the matrix T are given by the set

$$\bigcup_{\alpha \in \Lambda} \{\mu | \mu \text{ is eigenvalue of } \sum_{j=0}^{+\infty} Y_j \alpha^j\},$$

where Λ is the set of the eigenvalues of G; in particular, we deduce that the eigenvalues of the matrix R_F are eigenvalues of the matrix $\widehat{R}_F$. On the other hand, for the properties of nonnegative matrices [8], since for any $\alpha \in \Lambda$ it holds that

$$|\sum_{j=0}^{+\infty} Y_j \alpha^j| \leq \sum_{j=0}^{+\infty} Y_j |\alpha^j| \leq \sum_{j=0}^{+\infty} Y_j = R_F^T,$$

we find that $\rho(\widehat{R}_F) \leq \rho(R_F)$, whence $\rho(\widehat{R}_F) = \rho(R_F)$. From the properties of tensor product and from the relation $\mathbf{e}^T G = \mathbf{e}^T$, we obtain

$$\begin{array}{l} (\mathbf{y}^T \otimes \mathbf{e}^T) \widehat{R}_F = (\mathbf{y}^T \otimes \mathbf{e}^T) \sum_{j=0}^{+\infty} (Y_j \otimes G^j) = \\ \\ \sum_{j=0}^{+\infty} (\mathbf{y}^T Y_j) \otimes (\mathbf{e}^T G^j) = \sum_{j=0}^{+\infty} (\mathbf{y}^T Y_j) \otimes \mathbf{e}^T = \\ \\ \left(\mathbf{y}^T \sum_{i=0}^{+\infty} Y_j\right) \otimes \mathbf{e}^T = \lambda (\mathbf{y}^T \otimes \mathbf{e}^T). \end{array}$$

Therefore, the vector $\mathbf{w}^T = \mathbf{y}^T \otimes \mathbf{e}^T$ is a left nonnegative eigenvector of the matrix $\widehat{R}_F$. $\square$

By following the same argument of the proof of proposition 13, it is straightforward to prove the following

Proposition 14 *For every $n \geq 0$ the spectral radius of the matrix $\widehat{R}_n$ is given by*

$$\rho(\widehat{R}_n) = \rho\left(\sum_{j=0}^{+\infty} Y_{j,n}\right).$$

Moreover, a nonnegative k^2-dimensional vector $\mathbf{w}_n$ such that

$$\mathbf{w}_n^T \widehat{R}_n = \lambda_n \mathbf{w}_n^T,$$

where $\lambda_n = \rho(R_n)$, is given by $\mathbf{w}_n^T = (\mathbf{y}_n^T \otimes \mathbf{e}^T)$, where $\mathbf{y}_n$ is a nonnegative k-dimensional vector such that $\mathbf{y}_n^T \sum_{j=0}^{+\infty} Y_{j,n} = \lambda_n \mathbf{y}_n^T$.

From proposition 14 it follows that, for any integer $n \geq 0$, the left nonnegative eigenvector $\mathbf{w}_n$ of the matrix $\widehat{R}_n$ associated with its spectral radius λ_n belongs to the linear space $\mathcal{S}_1$ generated by the orthogonal vectors

$$\mathbf{v}_1 = \frac{1}{\sqrt{k}}(\mathbf{e}_1 \otimes \mathbf{e}), \quad \mathbf{v}_2 = \frac{1}{\sqrt{k}}(\mathbf{e}_2 \otimes \mathbf{e}), \ldots, \mathbf{v}_k = \frac{1}{\sqrt{k}}(\mathbf{e}_k \otimes \mathbf{e}), \tag{25}$$

where $\mathbf{e}_i$ is the k-dimensional vector having the i-th entry equal to 1 and the remaining entries equal to zero. Let $\mathcal{S}_2$ be the linear space $\mathcal{S}_2 = \mathcal{S}_1^{\perp}$. It can easily be observed that, since the matrices X_n and G are column stochastic, the error vector $\hat{\mathbf{e}}_n$ belongs to the linear space $\mathcal{S}_2$, for any n. In particular, if S_i is the matrix whose columns are an orthogonal basis of the linear space $\mathcal{S}_i$, for $i = 1, 2$, then the $k^2 \times k^2$ matrix $S = (S_1|S_2)$ is such that

$$S^T S = I, \quad S^T \hat{\mathbf{e}}_n = \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_n \end{pmatrix},$$

where $\mathbf{f}_n$ is a $k(k-1)$-dimensional vector. Whence the recursive equations (23) can be rewritten in the form

$$\begin{pmatrix} \mathbf{0} \\ \mathbf{f}_{n+1} \end{pmatrix} = (S^T \widehat{R}_n S)(S^T \hat{\mathbf{e}}_n) = (S^T \widehat{R}_n S) \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_n \end{pmatrix}. \tag{26}$$

On the other hand, observe that, if $\mathbf{w} \in \mathcal{S}_1$ then, from (24) and from the properties of tensor product, it follows that $(\mathbf{w}^T \widehat{R}_n)^T \in \mathcal{S}_1$, whence $S_1^T \widehat{R}_n S_2 = O$. Therefore the matrix $S^T \widehat{R}_n S$ has the structure

$$S^T \widehat{R}_n S = \begin{pmatrix} S_1^T \\ - \\ S_2^T \end{pmatrix} \widehat{R}_n \ (S_1|S_2) = \begin{pmatrix} V_n & O \\ T_n & W_n \end{pmatrix}, \tag{27}$$

where $V_n = S_1^T \widehat{R}_n S_1$, $W_n = S_2^T \widehat{R}_n S_2$ and $T_n = S_2^T \widehat{R}_n S_1$. Hence the recurrences (26) can be rewritten in the form

$$\mathbf{f}_{n+1} = W_n \mathbf{f}_n. \tag{28}$$

The above properties allow us to prove the following convergence results:

Theorem 15 *Let*

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|}$$

be the mean asymptotic rate of convergence of the sequence (22) where $F(\cdot)$ is defined by (15) and $\|\cdot\|$ is any vector norm. Then

$$r \le \rho(W_F)$$

where $W_F = S_2^T \widehat{R}_F S_2$ and $\widehat{R}_F = \lim_n \widehat{R}_n$.

Proof. From (26) and (28) it follows that

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|} = \lim_n \sqrt[n]{\|\mathbf{f}_n\|}.$$

Therefore, let us analyze the convergence of the sequence $\{\mathbf{f}_n\}$. Let $\epsilon > 0$ be fixed. Then (compare [10]) there exists a matrix norm $\|\cdot\|_\epsilon$ such that

$$\|W_F\|_\epsilon \le \rho(W_F) + \epsilon.$$

On the other hand, since the sequence $\{W_n\}$ converges to W_F, there exists an integer i_0 such that

$$\|W_i\|_\epsilon \le \|W_F\|_\epsilon + \epsilon$$

for any $i \ge i_0$. Therefore

$$\|\mathbf{f}_n\|_\epsilon = \|W_{n-1}W_{n-2}\dots W_0\mathbf{f}_0\|_\epsilon \le c_{i_0}\|W_{n-1}\|_\epsilon\|W_{n-2}\|_\epsilon \dots \|W_{i_0}\|_\epsilon \le$$

$$c_{i_0}(\|W_F\|_\epsilon + \epsilon)^{n-i_0} \le c_{i_0}(\rho(W_F) + 2\epsilon)^{n-i_0}$$

where c_{i_0} is a positive constant. Hence, by taking the n-th root, we have, for $n \to +\infty$,

$$r = \lim_n \sqrt[n]{\|\mathbf{f}_n\|_\epsilon} \le \rho(W_F) + 2\epsilon.$$

Therefore, from the arbitrarity of ϵ and from (29), we conclude the proof. $\square$

Observe that, from the structure (27) of $\widehat{R}_F$, it holds $\rho(W_F) \le \rho(\widehat{R}_F) = \rho(R_F)$. In the case where the matrix $\widehat{R}_F$ is irreducible the inequality is strict, as shown by the following:

Theorem 16 *If the matrix $\widehat{R}_F$ is irreducible, then the mean asymptotic rate of convergence*

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|}$$

of the sequence (22), where $F(\cdot)$ is defined by (15) and $\|\cdot\|$ is any vector norm, is

$$r \le \rho_2(\widehat{R}_F)$$

where $\rho_2(\widehat{R}_F)$ denotes the second largest modulus eigenvalue of the matrix $\widehat{R}_F$.

Proof. From theorem 15 it is sufficient to prove that

$$\rho(W_F) \le \rho_2(\widehat{R}_F).$$

First observe that, from (27), the eigenvalues of the matrix $\widehat{R}_n$ coincide with the eigenvalues of the matrices V_n and W_n. On the other hand, if the columns of the matrix S_1 are given by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of (25), then the matrix V_n is nonnegative and such that, for the properties of tensor product and from the structure (24) of $\widehat{R}_n$,

$$V_n = S_1^T \widehat{R}_n S_1 = \frac{1}{k}(I \otimes \mathbf{e}^T)\widehat{R}_n(I \otimes \mathbf{e}) = \sum_{j=0}^{+\infty} Y_{j,n}.$$

Therefore, for proposition 14, the spectral radius of the matrix V_n is equal to the spectral radius of the matrix $\widehat{R}_n$, for any n. On the other hand, since the matrix $\widehat{R}_F$ is irreducible, it follows that there exists an integer n_0 such that the matrices $\widehat{R}_n$ are irreducible for $n \ge n_0$. Whence, for the Perron Frobenius theorem (compare [2]), for any $n \ge n_0$, there exists a unique positive eigenvalue of the matrix $\widehat{R}_n$ having modulus equal to the spectral radius and its algebraic multiplicity is equal to 1. Therefore, for any $n \ge n_0$, the spectral radius of the matrix W_n is such that

$$\rho(W_n) \le \rho_2(\widehat{R}_n),$$

where $\rho_2(\widehat{R}_n)$ is the second largest modulus eigenvalue of the matrix $\widehat{R}_n$. In particular, by taking limits for $n \to +\infty$, we obtain

$$\rho(W_F) \le \rho_2(\widehat{R}_F), \tag{29}$$

where $W_F = \lim_n W_n = S_2^T \widehat{R}_F S_2$. □

Let us associate with a function $F(\cdot)$ of (15) the matrix

$$\begin{aligned}\widehat{R}_F = {} & \textstyle\sum_{i=1}^{+\infty} \left(G^{i-1}K_i(I - H(G))^{-1}\right)^T \otimes I + \\ & \textstyle\sum_{j=2}^{+\infty}\sum_{i=j}^{+\infty} \left(G^{i-j}A_i(I - H(G))^{-1}\right)^T \otimes G^{j-1} = \\ & \left((I - H(G))^{-1^T} \otimes I\right)\left(\textstyle\sum_{j=1}^{+\infty} \bar{A}_j^T \otimes G^{j-1} - H(G)^T \otimes I\right).\end{aligned}$$

From the above theorems and from the results of the previous section, it follows that, in the cases where the matrix $\widehat{R}_F$ is irreducible, then the mean asymptotic rate of convergence of the sequence $X_n = F(X_{n-1})$ is:

– equal to the spectral radius of $\widehat{R}_F$, if $X_0 = O$;

– less than or equal to the modulus of the second largest eigenvalue of $\widehat{R}_F$, if X_0 is a column stochastic matrix.

In many cases the modulus of the second largest eigenvalue of a matrix can be much smaller than its spectral radius. In the following figures we report, for the functional iterations defined by (4), (5), (6) the logarithm (to the base 10) of the residual $\|X_n - \sum_{i=0}^{+\infty} X_n^i A_i\|_1$, for the sequences obtained by starting with $X_0 = O$ and with $X_0 = I$, for a problem arising from the modelling of a metropolitan network [1]. It is worth pointing out the increasing of the speed of convergence obtained by starting with $X_0 = I$ instead of $X_0 = O$. Moreover it can be observed that the method defined by (6) (method based on the matrix U) converges more quickly than the methods based on (4) and (5) (natural and traditional algorithm, respectively). More precisely the spectral radius ρ_1 of the matrix $\widehat{R}_F$, which gives the mean asymptotic rate of convergence of the sequence obtained by starting with $X_0 = O$, is given by $\rho_1 = 0.998448$ for the natural algorithm defined by (4), $\rho_1 = 0.998379$ for the traditional algorithm defined by (5) and $\rho_1 = 0.998176$ for the algorithm based on the matrix U defined by (6). On the other hand the second largest modulus eigenvalue of the matrix $\widehat{R}_F$ is $\rho_2 = 0.883677$ for the natural algorithm, $\rho_2 = 0.875044$ for the traditional algorithm and $\rho_2 = 0.858737$ for the algorithm based on the matrix U. Moreover, as it can easily be observed from the figures, for this particular example the rate of convergence of the sequence obtained by starting with $X_0 = I$ is equal to the second largest modulus eigenvalue of the matrix $\widehat{R}_F$.

References

[1] Anastasi, G., Lenzini L. (1995): Personal communication.

[2] Berman A., Plemmons R.J. (1979): Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York.

[3] Bini D., Meini B. (1996): On the solution of a nonlinear matrix equation arising in queueing problems. *SIAM J. Matrix Anal. Appl.*, 17(4), pp. 906–926.

[4] Çinlar E. (1975): Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, N.J..

[5] Latouche G. (1992): Algorithms for evaluating the matrix G in Markov chains of PH/G/1 type. Bellcore Technical Report.

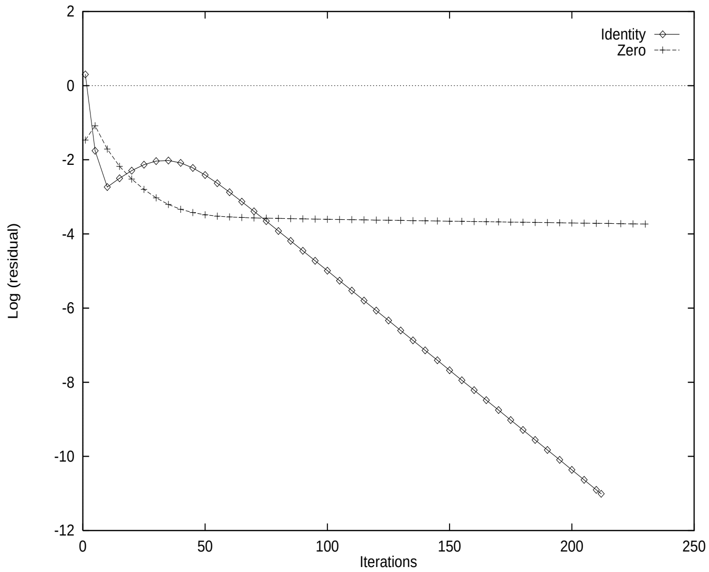

Figure 1: Natural algorithm

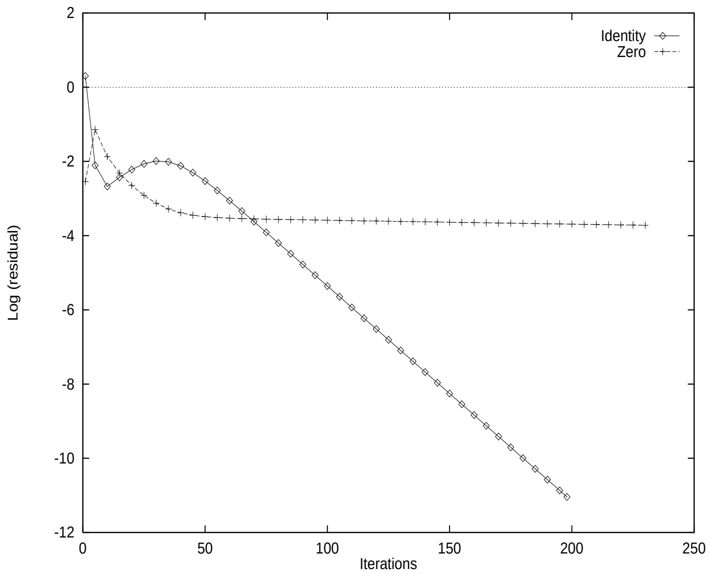

Figure 2: Traditional algorithm

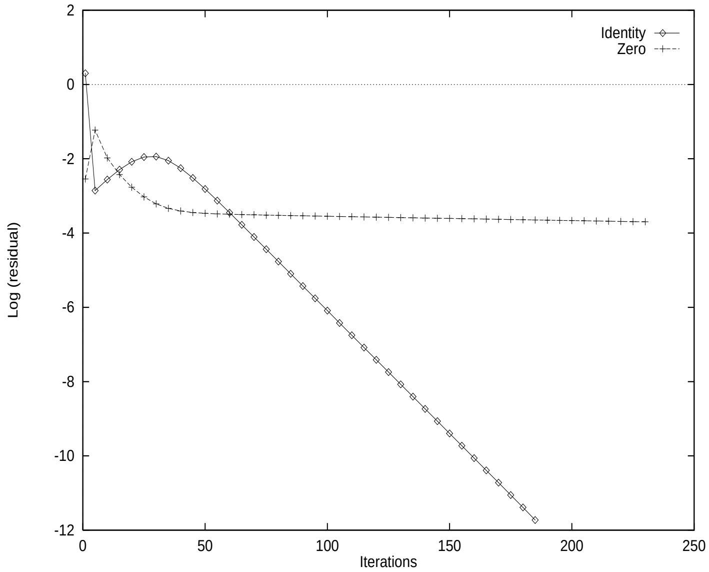

Figure 3: Algorithm based on matrix U

[6] Latouche G. (1995): Personal communication.

[7] Latouche G., Stewart G.W. (1996): Numerical methods for M/G/1 type queues. in Computations with Markov Chains, W. J. Stewart Ed., pp. 571–581, Kluwer Academic Publishers.

[8] Minc H. (1988): Nonnegative Matrices, John Wiley and Sons, Inc..

[9] Neuts M.F. (1989): Structured Stochastic Matrices of M/G/1 Type and Their Applications, Dekker Inc., New York.

[10] Varga R.S. (1962): Matrix Iterative Analysis, New Jersey, Prentice Hall .