Strong stability and perturbation bounds for discrete Markov chains

Boualem Rabta*, Djamil Aïssani

L.A.M.O.S., Laboratory of Modeling and Optimization of Systems, University of Bejaia, 06000, Algeria

Received 19 April 2007; accepted 28 October 2007

Abstract

We consider the stability problem of discrete Markov chains when their transition matrices are perturbed. For finite irreducible Markov chains many perturbation bounds for the stationary vector are available. In this paper, we identify a condition under which these bounds are carried over to discrete irreducible Markov chains which are not finite.

© 2007 Elsevier Inc. All rights reserved.

AMS classification: 60J10; 60J22

Keywords: Markov chain; Discrete state space; Perturbation; Strong stability; Quantitative estimates

1. Introduction

Since Schweitzer’s paper [15], several papers have been devoted to the investigation of the sensitivity of finite irreducible Markov chains to the perturbation of their transition matrices. The authors focused their efforts on the quantitative question, i.e., the obtention of an upper bound for the deviation of the stationary vector (or its individual components). While the fundamental matrix of Kemeny and Snell [10] is used by Schweitzer, Meyer [13] used the group inverse. This is also the case of Funderlic and Meyer [6], Haviv and Van der Heyden [7], Ipsen and Meyer [8] and Kirkland et al. [12]. Seneta [17,18] suggested the use of ergodicity coefficients and Cho and Meyer [3] expressed their bounds by means of mean first passage times. Eight perturbation

* Corresponding author. Tel.: +213 34 21 08 00; fax: +213 34 21 51 88.
E-mail addresses: brabta@yahoo.fr, lamos_bejaia@hotmail.com (B. Rabta).

0024-3795/S - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2007.10.036
bounds were collected and compared by Cho and Meyer in [4]. Most of them are given in terms of the $l_1$-norm (norm-wise bounds) and the absolute value (component-wise bounds) and this is why we will refer to this approach as the absolute stability method.

There is a significant body of literature on perturbation analysis for finite-state Markov chains and Markov processes. However, there are only few references available on perturbation analysis of Markov chains with an infinite state space [2]. Denumerable Markov chains can be used to represent many real systems. For example, the number of customers in a queue ($M/M/1$, $G/M/1$, $M/G/1$ ... ) at a given date is a random variable with an infinite state space. Therefore, it is important to have similar bounds valid for denumerable chains. In this paper, we identify a condition under which many results [15,13,7,6,17,18,12,3] previously obtained for the finite-state case are carried over to denumerable chains. Also, we show that this condition is equivalent to the strong stability of the underlying Markov chain.

Strong stable Markov chains were first studied by Aïssani and Kartashov [1]. Since, many conditions for the strong stability and many perturbation bounds have been established for Markov chains with general state space and with respect to a general class of norms (see, e.g., [9]). Definitely, those results are also valid for denumerable and finite chains. However, the consideration of the discrete case allows us to make several refinements and to obtain additional results as well. In particular, it will be possible to derive perturbation bounds for the individual components of the stationary vector (i.e., component-wise bounds).

The remaining of this paper is organized as follows: In Section 2, we introduce notations and preliminary results. The definition of the strong stability is given in Section 3. The main results are given in Section 4. First, we clarify the conditions for the strong stability of discrete Markov chains, then, we show the validity of several bounds. We terminate our paper by a conclusion.

2. Preliminary and notations

Let $X = (X_t)_{t \geq 0}$ be a discrete-time homogeneous Markov chain with discrete (finite or denumerable) state space $E$. Random transitions between states are given by the transition matrix $P = (p_{ij})_{i,j \in E}$ where, $p_{ij}$ is the probability to move from state $i$ to state $j$ in one step. Assume that the chain is irreducible and that it admits a unique stationary vector $\pi = (\pi_0, \pi_1, \pi_2, ...)^T$ such that
\[
\begin{align*}
\pi^T P &= \pi^T, \\
\pi^T e &= \sum_{i \in E} \pi_i = 1,
\end{align*}
\]
where $e = (1, 1, 1, ...)^T$ is the vector of all ones.

Suppose that the transition matrix $P$ of the chain $X$ is perturbed to $Q = P + \Delta$ such that $Q$ corresponds to the transition matrix of an irreducible Markov chain $X'$ on the same state space and with stationary vector $v = (v_0, v_1, v_2, ...)^T$.

In this paper, we use the norm $\| \cdot \|$ for row vectors and $\| \cdot \|_{\infty}$ for matrices and column vectors. Also, all vectors are column vectors. Row vectors are transposed, so it will be always clear if the norm corresponds to a matrix, a column vector or a row vector. All norms can be noted $\| \cdot \|$ without risk of confusion. Thus, for a vector $v$, we have
\[
\|v^T\| = \sum_{j \in E} |v_j| \quad \text{and} \quad \|v\| = \sup_{i \in E} |v_i|,
\]
and for a matrix $B = (b_{ij})_{i,j \in E}$, we have
∥B∥ = sup \sum_{i \in E} |b_{ij}|.

We notice that for every stochastic matrix \( P : ∥P∥ = 1 \) and for every probability vector \( \pi : ∥\pi^T∥ = 1 \).

Note that we will use sup and inf symbols, even for finite vectors and matrices, to avoid rewriting similar formulas.

Denote by \( A \) the matrix \( e\pi^T \) and by \( R = (I - P + II)^{-1} \) the fundamental matrix of the chain \( X \).

In 1958, Drazin [5] introduced a class of generalized inverses in a semi-group or an associative ring. A particular (Drazin) generalized inverse which plays an important role in matrix perturbation theory is the group inverse. For a square matrix \( A \), it is the matrix \( A^# \) satisfying

\[ AA^#A = A, A^#AA^# = A^# \]

If it exists, the Drazin-inverse of an element is unique [5].

In the sequel, \( A \) denotes the matrix \( I - P \). Hence, it is not hard to check that \( A^# = R - II \) so it becomes clear that the two quantities can be used in the same manner.

The matrix \( W = \sum_{i=0}^{\infty}(P^i - II) \) is called the deviation matrix of the chain \( X \). Many papers have been devoted to the investigation of the properties of \( W \) and the conditions for its existence (see, e.g. [11,19]). It is known that, when the matrix \( W \) exists, it is the group inverse of \( A \), i.e.

\[ A^# = W. \]

Let \( m_{ij} \), \( i, j \in E \), be the mean first passage time from state \( i \) to state \( j \). In [19, Corollary 3.3] Syski showed that when the deviation matrix \( W = (w_{ij})_{i,j \in E} \) exists it verifies:

\[ w_{ij} = w_{jj} - \pi_j m_{ij}, \quad i, j \in E, \]

Observe that \( IIW = 0 \), i.e., \( \sum_{i \in E} \pi_i w_{ij} = 0, \forall j \in E \).

Since \( W = A^# \) we may write:

\[ \pi_j m_{ij} = a^#_{jj} - a^#_{ij}, \quad i, j \in E, \]

Whence, for every \( k \in E \):

\[ \pi_k \sup_{i \in E} m_{ik} = \sup_{i \in E} (a^#_{kk} - a^#_{ik}). \]

Now, consider the ergodicity coefficient \( A_1(B) \) of a matrix \( B \) with equal row sums satisfying \( ∥B∥ < \infty \):

\[ A_1(B) = \sup\{∥u^T B∥ : u^T e = 0, ∥u^T∥ ≤ 1\}. \]

Having \( A^# = R - II \) and taking into account that \( u^T II = 0 \) for all \( u \) such that \( u^T e = 0 \), we deduce that for the Markov chain \( X \):

\[ A_1(A^#) = A_1(R). \]

It was mentioned in [14, p. 73] (see also [16]) that any real vector \( d = (d_i)_{i \in E} \) such that \( ∥d^T∥ < \infty \) and \( d^T e = 0 \), may be expressed in the form \( d = \sum_i \delta(i) \), where the vectors \( \delta(i) = (\delta_{ij})_{i,j \in E} \) have only two non-zero entries (of opposite sign), and \( ∥d^T∥ = \sum_i ∥\delta^T(i)∥ \). For a suitable set \( S = S(d) \) of ordered pairs of indices \((i, j)\), we may write [16],

\[ d = \sum_{(i,j) \in S} (\eta_{ij}/2)\gamma(i, j), \]
where $\eta_{ij}$ are positive scalars such that $\sum \eta_{ij} = \|d^T\|$, while the only non-zero entries of the vectors $\gamma(i, j)$ are $+1$ and $-1$, in the $i, j$ positions. We therefore have that, for each such $d$ and an arbitrary bounded vector $c = (c_i)_i \in E$:

$$d^T c = \sum_r d_r c_r = \sum_{(i, j) \in S} (\eta_{ij}/2)(c_i - c_j).$$

Hence

$$|d^T c| = \left| \sum_{(i, j) \in S} (\eta_{ij}/2)(c_i - c_j) \right| \leq \sum_{(i, j) \in S} (\eta_{ij}/2)|c_i - c_j|$$

(4)

$$|d^T c| \leq \frac{1}{2} \|d^T\| \sup_{i, j} |c_i - c_j|.$$

Also, Seneta [16] showed that

$$\|d^T P\| \leq \frac{1}{2} \sup_{i, j} \sum_s |p_{is} - p_{js}| \|d^T\|,$$

and used it to derive an explicit form of the ergodicity coefficient of a stochastic (finite or infinite) matrix. That is

$$A_1(P) = \frac{1}{2} \sup_{i, j} \sum_s |p_{is} - p_{js}|.$$

He noted that the argument is also valid for any matrix $B$ for which $\|d^T B\|$ is well defined. So, we observe that the ergodicity coefficient of a bounded infinite matrix with equal row sums is the same.

### 2.1. Perturbation bounds for finite Markov chains

The perturbation bounds we talk about in this paper are of two kinds. First, there are norm-wise bounds which are of the form:

$$\|v^T - \pi^T\| \leq B\|\Delta\|,$$

and then, component-wise perturbation bounds which are of the form:

$$|v_k - \pi_k| \leq b_k\|\Delta\|, k \in E,$$

where $B$ (and $b_k$) is a positive constant called the condition number of the chain. A bound of the second type provide us with a detailed information about the sensitivity of the chain to perturbations.

By observing that $\|(v - \pi)^T\| \leq 1$ and $(v - \pi)^T e = 0$ and by taking $c = e_k$, i.e., the $k$th column of the identity matrix, we can easily obtain from relation (4) that for every $k \in E$:

$$|v_k - \pi_k| \leq \frac{1}{2} \|v^T - \pi^T\|,$$

which allows us to deduce a component-wise bound from a norm-wise bound.

As we have already mentioned, there is a considerable amount of literature on perturbation analysis of finite irreducible Markov chains. In 1968, Schweitzer showed that [15]:

$$v^T - \pi^T = v^T \Delta R.$$
Later, Meyer [13] proved that the group inverse $A^#$ of $A = I - P$ can be used exactly in the same manner as $R$ and he gave the following result:

$$v^T - \pi^T = v^T \Delta A^#$$

which gives the bound:

$$\|v^T - \pi^T\| \leq \|\Delta\|A^\#\|.$$  \hspace{1cm} (6)

Also, Funderlic and Meyer [6] derived from (5) the bound:

$$|v_k - \pi_k| \leq \|\Delta\| \sup_{i \in E} |a^#_{ik}| \quad \forall k \in E.$$  \hspace{1cm} (7)

Relation (5) has also been used by Haviv and Van der Heyden in [7] and Kirkland et al. in [12] to show that

$$|v_k - \pi_k| \leq \frac{1}{2} \|\Delta\| \sup_{i \in E} (a^#_{kk} - a^#_{ik}) \quad \forall k \in E.$$  \hspace{1cm} (8)

In 2000, Cho and Meyer [3] suggested the use of mean first time passages and they gave the bound:

$$|v_k - \pi_k| \leq \frac{1}{2} \|\Delta\| \pi_k \sup_{i \in E} m_{ik} \quad \forall k \in E.$$  \hspace{1cm} (9)

If the ergodicity coefficient of the chain $X$ satisfies $A_1(P) < 1$, then we have the bound given by Seneta in [17]:

$$\|v^T - \pi^T\| \leq \frac{\|\Delta\|}{1 - A_1(P)}.$$  \hspace{1cm} (10)

Instead, if $A_1(P) = 1$, then we can use the following inequality [18]:

$$\|v^T - \pi^T\| \leq \|\Delta\| A_1(A^#).$$  \hspace{1cm} (11)

3. Strong stability

The strong stability method [1,9] considers the problem of the perturbation of general state space Markov chains using operators’ theory and with respect to a general class of norms. The basic idea behind the concept of stability is that, for a strongly stable Markov chain, a small perturbation in the transition matrix can lead to only a small deviation of the stationary vector. We will see in the next section that, for finite irreducible chains, this is always true but for denumerable (and general state space) chains, additional conditions are necessary.

**Definition 1.** A Markov chain $X$ with transition matrix $P$ and stationary vector $\pi$ is said to be strongly stable for the norm $\|\cdot\|$ if every stochastic matrix $Q$ in the neighbourhood ($\|Q - P\| \leq \varepsilon$ for some $\varepsilon > 0$) admits a unique stationary vector $v$ and:

$$\|v^T - \pi^T\| \to 0 \quad \text{as} \quad \|Q - P\| \to 0.$$  \hspace{1cm} (12)

In fact, as shown in [1], $X$ is strongly stable if and only if, there exists a positive constant $C = C(P)$ such that

$$\|v^T - \pi^T\| \leq C\|Q - P\|.$$  \hspace{1cm}

In another words, such a perturbation bound can be obtained only if the chain $X$ is stable. In case of non-stability, a small perturbation can lead to an important deviation. In fact, this deviation is always bounded (for the norm $\|\cdot\|_1$) since
\[ \|v^T - \pi^T\| \leq \|v^T\| + \|\pi^T\| = 2 \]

but, we may not have (12).

Note that some authors use the term “stability” in a different manner. For example, in [6], a chain \( X \) is said to be stable if the condition number is small where, the exact meaning of the term “small” has to be defined with respect to the underlying application.

4. Criteria for the strong stability of discrete Markov chains

The following theorem [1] gives a necessary and sufficient condition for the strong stability of Markov chains.

**Theorem 2.** The Markov chain \( X \) is strongly stable in the norm \( \|\cdot\| \) if and only if the matrix \( (I - P + \Pi) \) has a bounded inverse, i.e.

\[ \|(I - P + \Pi)^{-1}\| < \infty. \]

Note that for the denumerable case, Kemeny et al. [11] showed that the fundamental matrix of \( X \) exists under some particular conditions. Also, we know that \( R = A^\# + \Pi \). Then, we can use the group inverse \( A^\# \) in the same manner as \( R \).

The following is a sufficient (but not necessary) condition for the strong stability of \( X \).

**Corollary 3.** If the matrix \( W = \sum_{i=0}^{\infty} (P_i - \Pi) \) exists and is bounded then the Markov chain \( X \) is strongly stable.

**Proof.** This is straightforward from (1) and Theorem 2 (by using \( A^\# \) instead of \( R \)). □

For a finite irreducible Markov chain, the following result holds.

**Theorem 4.** A finite irreducible Markov chain is strongly stable (with respect to the norm \( \|\cdot\|_1 \)).

**Proof.** Kemeny and Snell [10] showed that for a finite irreducible Markov chain, the fundamental matrix exists. It is easy to verify that any finite matrix is bounded for the norm \( \|\cdot\|_1 \). □

The importance of this result is that it shows why no condition is required to the establishment of the above perturbation bounds in the finite case.

Under the condition of Theorem 2, relations (5), (6), (11), (8) and (7) hold for Markov chains with denumerable state space. The proofs are just the same as in the finite case.

Under the condition of Corollary 3, relation (9) follows straightforwardly from relation (8) together with (3).

The condition of Corollary 3 is required to prove relation (10). Furthermore, the comparison made by Cho and Meyer in [4] can be extended to the denumerable case with the same arguments.

5. Conclusion

In this paper, we gave a condition under which many perturbation bounds known for finite irreducible Markov chains, are carried over to the case where the state space is not necessarily finite. This condition, namely, the strong stability of the chain \( X \), is equivalent to the existence
and the boundedness of the fundamental matrix (or equivalently, the group inverse of $I - P$) of the chain.

References