Journal of Statistics Applications & Probability
An International Journal

Perturbation results for comparison of Markov models

Boualem Rabta

L.A.M.O.S., Laboratory of Modeling and Optimization of Systems, University of Bejaia, Targa Ouzemour, Algeria.
Email Address: brabta@yahoo.fr

Received: 10 Sep. 2012, Revised: 9 Nov. 2012, Accepted 19 Nov. 2012
Published online: 1 Mar. 2013

Abstract: Markov chains are useful to model various complex systems. In numerous situations, the underlying Markov chain is subject to changes. For example, states may be added or deleted and transition probabilities perturbed. It is therefore, necessary to ensure the robustness of the system and to estimate the resulting deviation in the characteristics. In this paper we study the sensitivity of finite Markov chains subject to changes in their state space and propose updating formulas and perturbation bounds.

Keywords: Markov chains, discrete state space, perturbation, strong stability, quantitative estimates, comparison.

1 Introduction

Markov chains are widely used in practice. Numerous real systems can directly or through a form of embedding, be described by Markov chains. However, a lot of real systems are very complicated and it might be necessary to use approximations and simplify the original model in order to make it tractable. Different kinds of perturbations can also affect the model (e.g., the parameters are generally estimated from empirical data by means of statistical methods). It is then necessary to study the robustness of the system and estimate the deviation in its characteristics due to all kinds of perturbations. In particular, the state space might be subject to changes, for example, when some states are added or deleted. Consider for instance, a simple queueing system with limited waiting capacity (or finite source). We might be interested to know how the system reacts when we increase or decrease the waiting capacity. Similarly, if we consider an (s, S) inventory system, then we may be interested in knowing the effect of different kinds of perturbations (see, e.g. [7]). In particular, the perturbation of the order-up-to level S. The on-hand inventory process X is a Markov chain (under some conditions) with states in $\{0, .., S\}$. Perturbation of the parameter S might be looked at as adding or deleting states. Another example is that of the Google search engine page rank. Google models the web as a huge Markov chain where web pages are states (nodes) and hyperlinks are transitions (see, e.g., [5]). Web pages are then ranked to allow ordering of search results. As there is a constant change in the internet where new pages are created and others are deleted every day, Google updates its web matrix and computes the updated page ranks in a constant basis. Therefore, it is important to derive updating formulas allowing the computation of the new characteristics from the old ones and perturbation bounds allowing measuring the difference between the original and the perturbed chains.

Updating formulas and perturbation bounds for Markov chain may be found in the literature. However, it is assumed that the changes affect the transition matrix of the chain, that is, the original and the perturbed Markov chains have the same state space. Schweitzer [9] used the fundamental matrix of Kemeny and Snell [4] to express the perturbation bound. Meyer [6] singly or with coauthors used the group inverse. Hunter [3] suggested the use of a wider class of generalized inverses and Seneta [10, 11] suggested the use of ergodicity coefficients. The survey paper by Cho and Meyer [2] collects and compares eight perturbation bounds. The qualitative question is treated in [8] where it is shown that a finite irreducible Markov chain is strongly stable. Strong stability [1] means that small changes in the inputs (transition matrix) can lead to only small deviations in the output (stationary vector and stationary characteristics). In other words, a strongly stable Markov chain reacts continuously to perturbations. In this paper, we focus on the problem of comparison of two Markov chains where their states spaces differ in some states. We provide updating formulas and perturbation bounds.

2 Preliminary and notations

Let $X = (X_n)_{n\geq 0}$ be a discrete-time homogeneous Markov chain with finite state space E_0. Random transitions between states are given by the transition matrix P_0. X admits a unique stationary vector $\pi^{(0)} = (\pi_0^{(0)}, \pi_1^{(0)}, \ldots, \pi_N^{(0)})^T$ satisfying :

$$\begin{cases} \pi^{(0)T}P_0 = \pi^{(0)T}, \\ \pi^{(0)T}e = \sum_{i\in E_0} \pi_i^{(0)} = 1. \end{cases}$$

Here and in the sequel, $e = (1,1,\ldots,1)^T$ is the vector of all ones of a suitable dimension. Also, $\mathbf{0}$ is either a vector or a matrix of all zeros and I is the identity matrix. Denote by Π_0 the matrix $e\pi^{(0)T}$ and let $Z_0 = (I - P_0 + \Pi_0)^{-1}$ be the fundamental matrix [4] of the chain X.
In this paper, we may use any matrix norm $\|.\|$ for which $\| P_0 \| < \infty$. A special interest is given to the norm $\|.\|_1$. Also, all vectors are column vectors. Row vectors are transposed.

3 Comparison of Markov chains

Now, we present a method to compare Markov chains with different state spaces. This problem might arise when some states are added (or removed) to the initial state space and such situations are multiple in practice.
Assume that the state space is changed by adding one or more states E_T and correcting the transition probabilities in such a way that the new Markov chain Y is irreducible with transition matrix Q and state space $E = E_0 \cup E_T$.
In order to be able to compare the two chains, we will first construct an intermediate Markov chain with the same characteristics as X and the same state space as Y. The transition matrix Q of the chain Y may be written in the form :

$$Q = \begin{pmatrix} Q_0 & Q_1 \\ Q_2 & Q_T \end{pmatrix}$$

where Q_0 and Q_T are square and they correspond respectively to states E_0 and E_T. Let $\widetilde{X}$ be the Markov chain with transition matrix :

$$\widetilde{P} = \begin{pmatrix} P_0 & \mathbf{0} \\ Q_2 & Q_T \end{pmatrix}.$$

Then, $\widetilde{X}$ has the stationary vector $\widetilde{\pi}^T = (\pi^{(0)T}, \mathbf{0}^T)$ and hence the stationary characteristics of the chain X. The idea is that we think about the newly added states as being transitive in the original state space and that the perturbation affects the transition matrix in such a manner that the resulting chain Y is irreducible. Let also Π^0 be the matrix $e\pi^{0T}$ and by $\widetilde{\Pi}$ the matrix $e\widetilde{\pi}^T$.

Remark 3.1 (States deletion) If we delete some states from state space (e.g. reducing waiting room capacity in a queueing system, decreasing the order-up-to level S in the (R, s, S) inventory system, etc.) then we can think of the deleted states as if they become transitive in the new chain. Hence, we keep the original Markov chain with transition matrix

$$P = \begin{pmatrix} P_0 & P_1 \\ P_2 & P_T \end{pmatrix}$$

and we construct a new intermediate Markov chain for the perturbed chain with transition matrix as follows

$$\tilde{Q} = \begin{pmatrix} Q & \mathbf{0} \\ P_2 & P_T \end{pmatrix}.$$

It appears that we can restrict our analysis to only one case (either adding or deleting state) because of symmetry.

The first step consists in checking the qualitative property of robustness of the Markov chain X. One has to be sure that the chain reacts continuously to the perturbations so that small deviations in the input (the transition matrix) will result in a bounded deviation of the output (stationary vector). Therefore, we prove that

Theorem 3.1 *The Markov chain $\tilde{X}$ is strongly stable with respect to the norm $\|.\|_1$, i.e., for each irreducible stochastic matrix Q in the neighborhood, having a stationary vector ν we have :*

$$\| \nu^T - \tilde{\pi}^T \|_1 \to 0 \quad as \quad \| Q - \tilde{P} \|_1 \to 0.$$

Proof. Let Q be an irreducible stochastic matrix with stationary vector ν. We may write :

$$\nu^T - \tilde{\pi}^T = \nu^T Q - \tilde{\pi}^T \tilde{P}$$

$$\nu^T - \tilde{\pi}^T = \nu^T (Q - \tilde{P}) + (\nu^T - \tilde{\pi}^T)\tilde{P}$$

$$(\nu^T - \tilde{\pi}^T)(I - \tilde{P}) = \nu^T (Q - \tilde{P}).$$

Since $(\nu^T - \tilde{\pi}^T)\tilde{\Pi} = 0$, we have :

$$(\nu^T - \tilde{\pi}^T)(I - \tilde{P} + \tilde{\Pi}) = \nu^T (Q - \tilde{P})$$

Since E_T is transitive, $Q_T^n \to 0$ as $n \to 0$ implying that $\exists n_0 > 0$ such that $\| Q_T^{n_0} \| < 1$. This means that $(I - Q_T)$ is invertible with inverse $(I - Q_T)^{-1} = \sum_{i=0}^{\infty} Q_T^i$. Also, the matrix

$$(I - \tilde{P} + \tilde{\Pi}) = \begin{pmatrix} I - P_0 + \Pi_0 & \mathbf{0} \\ e\pi^{(0)T} - Q_2 & (I - Q_T) \end{pmatrix}$$

is invertible with inverse

$$\tilde{Z} = \begin{pmatrix} Z_0 & \mathbf{0} \\ -(I - Q_T)^{-1}(e\pi^{(0)T} - Q_2)Z_0 & (I - Q_T)^{-1} \end{pmatrix}$$

Then,

$$\nu^T - \tilde{\pi}^T = \nu^T (Q - \tilde{P})\tilde{Z}. \tag{1}$$

and

$$\| \nu^T - \tilde{\pi}^T \|_1 \leq \| \nu^T \|_1 \| Q - \tilde{P} \|_1 \| \tilde{Z} \|_1$$

$$\| \nu^T - \tilde{\pi}^T \|_1 \leq C \| Q - \tilde{P} \|_1$$

where C is constant. Whence, the strong stability of the chain $\tilde{X}$.

Let

$$B = Q_1 (I - Q_T)^{-1}. \tag{2}$$

and

$$A = \left[(Q_0 - P_0) - B(e\pi^{(0)T} - Q_2)\right] Z_0 \tag{3}$$

In the sequel, we suppose that the following condition holds :

$$\exists m > 0 \quad \text{such that} \quad \rho = \| A^m \| < 1. \tag{4}$$

This means that $(I - A)$ is invertible with inverse $(I - A)^{-1} = \sum_{i=0}^{\infty} A^i$. Furthermore,

$$\| (I - A)^{-1} \| = \| \sum_{i=0}^{\infty} A^i \| \leq \sum_{i=0}^{\infty} \| A^i \|$$

by writing i as $i = km + r$, we obtain

$$\| (I - A)^{-1} \| \leq \sum_{k=0}^{\infty} \sum_{r=0}^{m-1} \| A^r A^{km} \| \leq \sum_{k=0}^{\infty} \left(\sum_{r=0}^{m-1} \| A^r \| \right) \| A^m \|^k$$

$$\| (I-A)^{-1} \| \leq \tau \sum_{k=0}^{\infty} \rho^k = \frac{\tau}{1-\rho}$$

where,

$$\tau = \sum_{r=0}^{m-1} \| A^r \| \leq m \max_{\{0 \leq r \leq m-1\}} \| A^r \|. \quad (5)$$

It is possible to compute the stationary vector of the perturbed chain directly from that of the original chain. The following result gives an updating formula for the stationary vector.

Theorem 3.2 *The stationary vector ν of the chain Y is given by :*

$$\nu^T = \pi^{(0)T}(I-A)^{-1}(I,B)$$

where A and B are given by (3) and (2) respectively.

Proof. From equation (1),

$$\nu^T = \tilde{\pi}^T (I-(Q-\tilde{P})\tilde{Z})^{-1}$$

where, $\tilde{Z} = (I - \tilde{P} + \tilde{\Pi})^{-1}$.

$$(Q-\tilde{P})\tilde{Z} = \begin{pmatrix} (Q_0 - P_0) & Q_1 \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} Z_0 & \mathbf{0} \\ -(I-Q_T)^{-1}(e\pi^{(0)T} - Q_2)Z_0 & (I-Q_T)^{-1} \end{pmatrix}$$

$$(Q-\tilde{P})\tilde{Z} = \begin{pmatrix} A & B \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

with,

$$A = \left[(Q_0 - P_0) - Q_1(I-Q_T)^{-1}(e\pi^{(0)T} - Q_2)\right]Z_0$$

and

$$B = Q_1(I-Q_T)^{-1}.$$

Thus,

$$(I-(Q-\tilde{P})\tilde{Z})^{-1} = \begin{pmatrix} (I-A)^{-1} & ((I-A)^{-1}B) \\ \mathbf{0} & I \end{pmatrix}$$

Finally,

$$\nu^T = (\pi^{(0)T}, \mathbf{0}^T) \begin{pmatrix} (I-A)^{-1} & ((I-A)^{-1}B) \\ \mathbf{0} & I \end{pmatrix}$$

$$\nu^T = (\pi^{(0)T}(I-A)^{-1}, \pi^{(0)T}(I-A)^{-1}B).$$

The updating formula given by Theorem 3.2, allows us to compute the stationary vector of the chain Y from the stationary vector of the chain X.

An important question in sensitivity analysis is the estimation of the difference between the characteristics of the original and the perturbed systems. This is in particular very important if one model is to be used as approximation to the other. Thus, if we need to estimate the changes made to the stationary probabilities, then we may use the following results.

Theorem 3.3 *The difference between the stationary vectors of the chains $\tilde{X}$ and Y is given by :*

$$\| \nu^T - \tilde{\pi}^T \|_1 \leq \| A \|_1 + \| B \|_1.$$

where A and B are given by (3) and (2) respectively.

Proof. From equation (1) we have :

$$\nu^T - \tilde{\pi}^T = \nu^T (Q-\tilde{P})\tilde{Z}$$

$$\| \nu^T - \tilde{\pi}^T \|_1 \leq \| \nu^T \|_1 . \| (Q-\tilde{P})\tilde{Z} \|_1$$

where $\| \nu^T \|_1 = 1$ and

$$\| (Q-\tilde{P})\tilde{Z} \|_1 = \left\| \begin{pmatrix} A & B \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \right\|_1 \leq \| A \|_1 + \| B \|_1.$$

Observe that the use of the norm $\|.\|_1$ allows us to obtain a perturbation bound without the need to compute the stationary vector of the perturbed chain Y since $\| \nu^T \|_1 = 1$ for any probability vector ν^T.

To allow the use of a more general class of norms we present the following result.

Theorem 3.4 *The difference between the stationary vectors of the chains $\tilde{X}$ and Y is given by :*

$$\| \nu^T - \tilde{\pi}^T \| \leq \| \pi^{(0)T} \| . \| (A, B) \| \frac{\tau}{1-\rho}.$$

where A and B are given by (3) and (2) respectively, ρ is defined in (4) and τ in (5).

Proof. The stationary vector ν of the chain Y is given by :

$$\nu^T = (\pi^{(0)T}(I-A)^{-1}, \pi^{(0)T}(I-A)^{-1}B).$$

Then,

$$\nu^T - \tilde{\pi}^T = (\pi^{(0)T}(I-A)^{-1} - \pi^{(0)T}, \pi^{(0)T}(I-A)^{-1}B)$$

$$\nu^T - \tilde{\pi}^T = (\pi^{(0)T}(I-A)^{-1}[I-(I-A)], \pi^{(0)T}(I-A)^{-1}B)$$

$$\nu^T - \tilde{\pi}^T = \pi^{(0)T}(I-A)^{-1}(A, B).$$

Thus,

$$\| \nu^T - \tilde{\pi}^T \| = \| (\pi^{(0)T}(I-A)^{-1}(A, B) \|$$

$$\| \nu^T - \tilde{\pi}^T \| \leq \| \pi^{(0)T} \| . \| (I-A)^{-1} \| . \| (A, B) \|$$

$$\| \nu^T - \tilde{\pi}^T \| \leq \| \pi^{(0)T} \| . \| (A, B) \| \frac{\tau}{1-\rho}.$$

Finally, it is worthy to note that the matrix $\tilde{A}^{\#} = \tilde{Z} - \tilde{\Pi}$ is the group generalized inverse [6] of the matrix $(I - \tilde{P})$ and may be used in the same manner as $\tilde{Z}$ (other generalized inverses [3] may also be used). Furthermore, similar perturbation bounds of those cited in [2, 8] may be derived in the same manner as in the classical results of perturbation theory.

References

[1] Aïssani, D. and Kartashov, N.V. Ergodicity and stability of Markov chains with respect to operator topology in the space of transition kernels. *Dokl. Akad. Nauk. Ukr. SSR*, ser. A 11, 3–5, (1983).

[2] Cho, G.E. and Meyer, C.D. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335, 137–150, (2001).

[3] Hunter, J.J. Generalized inverses and their applications to applied probability problems. *Linear Algebra and its Applications*, 45, 157–198, (1982).

[4] Kemeny, J.G. and Snell, J.L. *Finite Markov chains*, Van Nostrand, New York, (1960).

[5] Langville, A.N. and Meyer, C. D. Deeper Inside PageRank, *Internet Mathematics*, 1, 335–380, (2005).

[6] Meyer, C.D. The condition of a finite Markov chain and perturbation bounds for the limiting probabilities. *SIAM J. Algebraic Discrete Methods*, 1, 273–283, (1980).

[7] Rabta, B. and Aïssani, D. Strong Stability in an (R, s, S) Inventory Model. *International Journal of Production Economics*, 97, 159–171, (2005).

[8] Rabta, B. and Aïssani, D. Strong Stability and Perturbation Bounds for Discrete Markov Chains. *Linear Algebra and its Applications*, 428, 1921–1927, (2008).

[9] Schweitzer, P. Perturbation theory and finite Markov chains. *Journal of Applied Probability*, 5, 410–413, (1968).

[10] Seneta, S. Perturbation of the stationary distribution measured by ergodicity coefficients. *Advances in Applied Probability*, 20, 228–230, (1988)..

[11] Seneta, E. Sensitivity analysis, ergodicity coefficients, and rank-one updates for finite Markov chains. In W.J. Stewart, editor, Numerical solution of Markov chains, volume 8 of *Probab. Pure Appl.*, pages 121–129 (Marcel Dekker, New York), (1991).