Martin Kocurek
Insensitivity analysis of Markov chains

In: Jan Chleboun and Petr Přikryl and Karel Segeth and Jakub Šístek (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Dolní Maxov, June 6-11, 2010. Institute of Mathematics AS CR, Prague, 2010. pp. 107–112.

Persistent URL: http://dml.cz/dmlcz/702747

INSENSITIVITY ANALYSIS OF MARKOV CHAINS*

Martin Kocurek

Abstract

Sensitivity analysis of irreducible Markov chains considers an original Markov chain with transition probability matix P and modified Markov chain with transition probability matrix $\tilde{P}$. For their respective stationary probability vectors $\pi, \tilde{\pi}$, some of the following charactristics are usually studied: $\|\pi - \tilde{\pi}\|_p$ for asymptotical stability [3], $|\pi_i - \tilde{\pi}_i|$, $\frac{|\pi_i - \tilde{\pi}_i|}{\pi_i}$ for componentwise stability or sensitivity[1]. For functional transition probabilities, $P = P(t)$ and stationary probability vector $\pi(t)$, derivatives are also used for studying sensitivity of some components of stationary distribution with respect to modifications of P [2].

In special cases, modifications of matrix P leave certain stationary probabilities unchanged. This paper studies some special cases which lead to this behavior of stationary probabilities.

1 Introduction

A Markov chain is a sequence of random variables $X_1, X_2, X_3, \dots,$ with the Markov property, namely that, given the present state, the future and past states are independent. Formally,

$$P(X_{n+1} = x | X_1 = x_1, X_2 = x_2 \dots, X_n = x_n) = P(X_{n+1} = x | X_n = x_n),$$

where the possible values of X_i form a countable state space S of the chain. Markov chains are often described by a directed graph, where the edges are labeled by the probabilities p_{ij} of moving from state i to the other state j. These probabilities are called *transition probabilities* and together they form a *transition probability matrix* denoted by P, with row sums equal to 1. We will study finite Markov chains (a finite chain has a finite state space $S = \{x_1, ..., x_n\}$). A state i has period p if any return to state i must occur in multiples of p time steps. Formally, the period of a state i is defined as $p = gcd\{k : P(X_k = i | X_0 = i) > 0\}$. If $p = 1$, then the state is said to be aperiodic i.e. returns to state i can occur at irregular times. Otherwise ($p > 1$), the state is said to be periodic with period p. If all states are periodic with period p, the chain is called p-cyclic.

Let us denote

$$e = (1, \dots, 1)^T,\ e_i = (0, \dots, 0, 1, 0, \dots, 0)^T = (\delta_{i,j})_{j=1}^n,\ i = 1, \dots, n,\ P = (P_{ij})_{i,j=1}^n.$$

A Markov chain is called irreducible, if there exists a connection between every two states. That means, matrix P is irreducible. In this case, matrix P has a unique

*This work was supported by grant OHK1-062/10 of the Czech Technical University.

eigenvalue 1 (which equals to spectral radius $\rho(P)$ of P) and unique left and right eigenvectors associated with this eigenvalue, $\pi = (\pi_1, \dots, \pi_n)$ and e, so that

$$\pi P = \pi, \;\; Pe = e.$$

Vector π is called *stationary probability vector*, we usually normalise this vector to $\pi e = \|\pi\|_1 = 1$; i-th component π_i of π shows, how often the chain "visits state i",

$$\pi_i = \lim_{m\to\infty} \frac{|\{j; X_j = x_i, j = 1, \dots, m\}|}{m}.$$

We will also use a different normalisation, $\pi_k = 1$ and in this case, the eigenvector will be denoted by $\pi_{(k)}$, so that $\pi_{(k)k} = 1$.

In the following, we will partition matrix P and vector π into subblocks,

$$\pi = (\pi^{(1)}, \dots, \pi^{(N)}), \quad P = \begin{pmatrix} P_{11} & \dots & P_{1N} \\ \vdots & \ddots & \vdots \\ P_{N1} & \dots & P_{NN} \end{pmatrix}, \tag{1}$$

where N is the number of subblocks in matrix P, $n_1, ..., n_N$ will be respective dimensions of subblocks. Conformally with partitioning of P we shall partition vector $e = (e^{(1)T}, \dots, e^{(N)T})^T$, where $e^{(i)}$ is a vector $(1, \dots, 1)^T$ with n_i components.

As an example we will use a Markov chain with the following matrix:

$$P_c = \frac{1}{64} \begin{pmatrix}
62 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
62 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 60 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
62 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 63 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
62 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 60 & 2 & 0 & 0 & 0 & 0 & 0 \\
62 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 63 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 62 & 0 & 2 & 0 & 0 \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 60 & 0 & 0 \\
0 & 0 & 0 & 0 & 62 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 60
\end{pmatrix}. \tag{2}$$

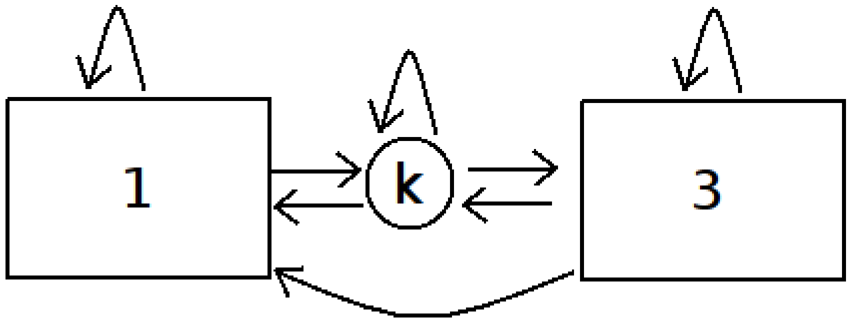
2 Normalisation $\pi_k = 1$

Normalisation $\pi_k = 1$ is useful for computing the eigenvector as a solution of a system of equations $\pi_{(k)}P = \pi_{(k)}$, or $P^T\pi_{(k)}^T = \pi_{(k)}^T$, $(I - P^T)\pi_{(k)}^T = 0$. By replacing an arbitrary equation with equation $\pi_{(k)}e_k = 1$, or equivalently, $e_k^T\pi_{(k)}^T = 1$, we obtain a system with better spectral properties, than when using condition $e^T\pi^T = 1$ [4].

When we use this normalization, we can state the following simple theorem.

Theorem 1. *Let the state space of a Markov chain can be decomposed into three groups* $S_1, \{x_k\} = S_2, S_3$, *so that in oriented graph of the Markov chain each path from* S_1 *to* S_3 *contains a vertex* x_k. *Then no modifications of transition probabilities between states of* S_1 *affect components in* $\pi_{(k)}$ *associated with states from* S_3

Proof: With given restrictions, the graph of the chain can be simplified into



At this picture, S_1 is denoted by 1, x_k by k, S_3 by 3. It then follows that nonzero structure of P is

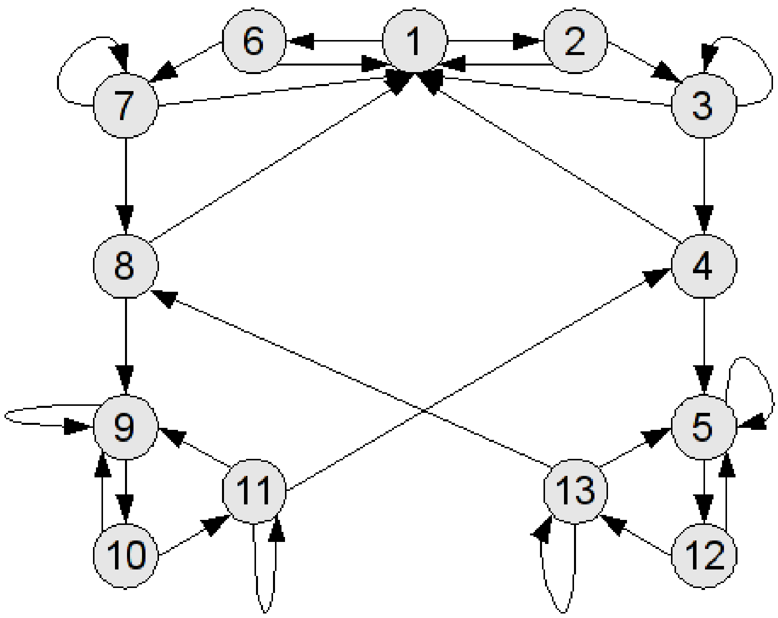
$$P = \left(\begin{array}{ccc|c|ccc} X & \dots & X & X & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X & \dots & X & X & 0 & \dots & 0 \\ \hline X & \dots & X & X & X & \dots & X \\ \hline X & \dots & X & X & X & \dots & X \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X & \dots & X & X & X & \dots & X \end{array}\right).$$

After forming left-hand side matrix $(I - P^T)$, we remove k-th equation and replace it with $e_k^T \pi_{(k)}^T = 1$. This way we obtain a system of equations with matrix $A^{(k)}$ and right-hand side e_k. $A^{(k)}$ has the following nonzero structure

$$A^{(k)} = \left(\begin{array}{ccc|c|ccc} X & \dots & X & X & X & \dots & X \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ X & \dots & X & X & X & \dots & X \\ \hline 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & X & X & \dots & X \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & X & X & \dots & X \end{array}\right),$$

it is clearly reducible. Thus, no modifications of transition probabilities between states in S_1 (block $1,1$ in $A^{(k)}$) will affect k-th,$\dots,n$-th components of $\pi_{(k)}$

Example: In example with P_c, we may draw an oriented graph:



We see that vertices $2, 3, 6, 7$ are accesible only through vertex 1. Thus if we fix the first element of π, no modifications of transition probabilities between vertices 4, 5, 8, 9, 10, 11, 12, 13 will affect components no. $2, 3, 6, 7$ in $\pi_{(1)}$.

3 Normalisation $\pi e = 1$

For a more usual normalisation $\pi e = 1$, let us first introduce a concept of lumpability

Definition: *Let us partition a transition probability matrix P into blocks $(P_{ij})_{i,j=1}^{N}$ so that for every block P_{ij} and vector $e^{(j)}$ of appropriate dimensions*

$$P_{ij}e^{(j)} = \alpha_{ij}e^{(j)}$$

for some $\alpha_{ij} \in \mathbf{R}$. Then matrix P is said to be lumpable.

Theorem 2. *Let $P(t)$ be a perturbed transition probability matrix of an irreducible finite aperiodic Markov chain, whose state space divided into subsets $S_1, \ldots, S_{N+1}$, where states of S_{N+1} are accessible only through S_N. Let perturbations depend on a variable t and be restricted to lumpable submatrix of blocks $(P_{ij}(t))_{i,j=1}^{N-1}$. If for every $i = 1, \ldots, N-1$ exists a column vector $x^{(i)}$ such that*

$$P_{i,N} = e^{(i)} \cdot x^{(i)T}, \tag{3}$$

then subblocks $\pi^{(N)}, \pi^{(N+1)}$ are independent of t

Proof: From the assumption it follows that

$$P_{ij}e^{(j)} = \alpha_{ij}e^{(j)},\ P_{i,N+1} = 0,\ i,j = 1,\ldots,N-1. \tag{4}$$

We will prove the theorem by using a power method for computing π. Assumptions guarrantee the existence of a unique steady point – eigenvector π [4].

Let us choose a $\pi^{(0)} = \left(\pi_1^{(0)},\ldots,\pi_{N+1}^{(0)}\right)$, for $l = 1,2,\ldots$

$$\pi^{(l+1)} = \pi^{(l)}P.$$

a) At first we will show by induction, that for every l the $\|\cdot\|_1$-norms of subvectors $\pi_1^{(l)},\ldots,\pi_{N+1}^{(l)}$ of $\pi^{(l)}$ do not depend on t. $\pi^{(0)}$ does not depend on t. The l_1-norm of the j-th subvector, $j = 1,\ldots,N-1$, in the $(l+1)$-th iteration is

$$\|\pi_j^{(l+1)}\|_1 = \pi^{(l)}P_{*,j}e^{(j)} = \sum_{i=1}^{N-1}\pi_i^{(l)}P_{i,j}(t)e^{(j)} + \sum_{i=N}^{N+1}\pi_i^{(l)}P_{i,j}e^{(j)} =$$

$$\sum_{i=1}^{N-1}\pi_i^{(l)}\alpha_{i,j}e^{(j)} + \sum_{i=N}^{N+1}\pi_i^{(l)}P_{i,j}e^{(j)},$$

which does not depend on t. For $j = N, N+1$ subblocks $P_{*,j}$ do not depend on t, thus $\|\pi_j^{(l+1)}\|_1 = \pi^{(l)}P_{*,j}e^{(j)}$ is also independent of t.

b) Now let us suppose that in iteration $\pi^{(l)}$ subvectors $N, N+1$ are independent of t. First, by (4) we have

$$\pi_{N+1}^{(l+1)} = \sum_{i=1}^{N+1}\pi_i^{(l)}P_{i,N+1} = \sum_{i=N}^{N+1}\pi_i^{(l)}P_{i,N+1},$$

which by induction hypothesis does not depend on t.

Finally, because of (3),

$$\pi_N^{(l+1)} = \sum_{i=1}^{N+1}\pi_i^{(l)}P_{i,N} = \sum_{i=1}^{N-1}\pi_i^{(l)}e^{(i)}x^{(i)T} + \pi_N^{(l)}P_{N,N} + \pi_{N+1}^{(l)}P_{N+1,N} =$$

$$= \sum_{i=1}^{N-1}\|\pi_i^{(l)}\|_1 x^{(i)T} + \pi_N^{(l)}P_{N,N} + \pi_{N+1}^{(l)}P_{N+1,N},$$

with all terms independent of t.

Remark: The above theorem holds also for periodic chains. If P is a transition probability matrix of p-cyclic chain, it has exactly p eigenvalues on a unit circle (one of them being 1). If we transform matrix P onto

$$\tilde{P} = \alpha P + (1-\alpha)I,$$

we obtain a matrix with submatrix of subblocks $(\tilde{P}_{ij}(t))_{i,j=1}^{N-1}$ remaining lumpable and for $i = 1, \ldots, N-1$ we will have $P_{i,N} = e^{(i)} \cdot \alpha x^{(i)T}$. Furthermore $\pi\tilde{P} = \pi$ and all eigenvalues other than 1 will be inside the unit circle, ensuring convergence of power method.

Example: If we change the order of states in Markov chain represented by P_c to 13, 11, 12, 10, 5, 9, 4, 8, 1, 2, 6, 3, 7, then the resulting chain has a transition probability matrix (zeros omited)

$$\bar{P}_c = \frac{1}{64}\left(\begin{array}{cc|cc|cc|cc|c|cc|cc}
60 & & & & 2 & & & 2 & & & & & \\
 & 60 & & & & 2 & 2 & & & & & & \\
\hline
2 & & & & 62 & & & & & & & & \\
 & 2 & & & & 62 & & & & & & & \\
\hline
 & & 1 & & 63 & & & & & & & & \\
 & & & 1 & & 63 & & & & & & & \\
\hline
 & & & & 2 & & & & 62 & & & & \\
 & & & & & 2 & & & 62 & & & & \\
\hline
 & & & & & & & & 62 & 1 & 1 & & \\
\hline
 & & & & & & & & 62 & & & 2 & \\
 & & & & & & & & 62 & & & & 2 \\
\hline
 & & & & & & 2 & & 2 & & & 60 & \\
 & & & & & & & 2 & 2 & & & & 60
\end{array}\right),$$

which is lumpable and if we have perturbations for example

$$\bar{p}_{11}(t) = \frac{60}{64} - t,\ \bar{p}_{12}(t) = t,\ \bar{p}_{66(t)} = \frac{63}{64} - 2t,\ \bar{p}_{65}(t) = 2t,$$

the resulting matrix satisfies conditions of the theorem.

4 Summary

This paper intends to present some conditions for insensitivity of a Markov chain towards perturbations in transition probability matrix. These conditions involve existence of cutpoints and regularity described by the concept of lumpability.

References

[1] Cho, G.E. and Meyer, C.D.: Markov chain sensibility measured by mean first passage times. Lin. Alg. Appl. **316** (2000), 21–28.

[2] Deutch, E. and Neumann, M.: On the derivatives of the Perron vector. Portugaliae Mathematica **43** (1985-1986), 35–42.

[3] Kirkland, S.J., Neumann, M. and Sze, N.-S.: On optimal condition numbers for Markov chains. Numerische Mathematik **110** (2008), 521–537.

[4] Stewart, W.J.: *Introduction to the numerical solution of Markov chains.* Princeton University Press, 1994.