Journal of Applied Analysis
Vol. 4, No. 2 (1998), pp. 259–267

UNCOUPLING MEASURES AND EIGENVALUES OF STOCHASTIC MATRICES

P. POKAROWSKI

Received June 15, 1998 and, in revised form, September 28, 1998

Abstract. This paper gives bounds for the uncoupling measures of a stochastic matrix $\mathbf{P}$ in terms of its eigenvalues. The proofs are combinatorial. We use the Matrix–Tree Theorem which represents principal minors of $\mathbf{I}-\mathbf{P}$ as sums of weights of directed forests.

1. Introduction and main results

Let $S := \{1, 2, \dots, n\}$ and $\mathbf{P} = (p_{ij})_{i,j=1}^{n}$ be an $n \times n$ stochastic matrix. Define the uncoupling measures of $\mathbf{P}$ (see [4]) as

$$\sigma_k := \min_{\substack{\emptyset \neq C_1,\dots,C_k \subset S \\ C_i \cap C_j = \emptyset \\ i,j=1,\dots,k}} \sum_{l=1}^{k} \sum_{\substack{i \in C_l \\ j \notin C_l}} p_{ij}, \quad k = 2, \dots, n,$$

1991 *Mathematics Subject Classification.* 15A42, 15A51, 60J10, 65F10, 65F15.

Key words and phrases. Stochastic matrices, nearly completely decomposable Markov chains, eigenvalues, uncoupling measures, directed forests, Matrix–Tree Theorem, numerical methods for linear systems.

ISSN 1425-6908

where $A \subset B$ means that $A \subseteq B$ and $A \neq B$. We will also investigate the following new measures of uncoupling:

$$\mu_k := \min_{\substack{\emptyset \neq C_1,\dots,C_k \subset S \\ C_i \cap C_j = \emptyset \\ i,j=1,\dots,k}} \max_{1 \le l \le k} \max_{\substack{i \in C_l \\ j \notin C_l}} p_{ij}, \quad k = 2,\dots,n.$$

The first proposition is a direct consequence of the above definitions.

Proposition 1. *For* $k = 2,\dots,n-1$

(i) $\sigma_k \le \sigma_{k+1}$;
(ii) $\mu_k \le \mu_{k+1}$;
(iii) $\mu_k \le \sigma_k$.

It is easy to check by induction on k that also the following complementary inequality holds.

Proposition 2. *For* $k = 2,\dots,n-1$

$$\sigma_k \le \frac{k-1}{k} n^2 \mu_k.$$

Let $\mathbf{I}$ be an $n \times n$ identity matrix and $\lambda_1, \lambda_2, \dots, \lambda_n$ be eigenvalues of $\mathbf{I} - \mathbf{P}$ indexed in such a way that

$$\lambda_1 = 0 \le |\lambda_2| \le \ldots \le |\lambda_n|.$$

Obviously λ is an eigenvalue of $\mathbf{I} - \mathbf{P}$ if and only if $1 - \lambda$ is an eigenvalue of $\mathbf{P}$. Our main results are bounds between uncoupling measures σ_k and eigenvalues of $\mathbf{I} - \mathbf{P}$ and, what is the same, eigenvalues of $\mathbf{P}$.

Theorem 1. *For* $k = 2,\dots,n$

$$\sigma_k \ge \left[\binom{n-1}{k-1}^{-1} |\lambda_2| \cdots |\lambda_k| \right]^{1/(k-1)}.$$

In particular we have

$$\sigma_2 \ge \frac{1}{n-1} |\lambda_2|,$$

$$\sigma_3 \ge \left[\frac{2}{(n-1)(n-2)} |\lambda_2 \lambda_3| \right]^{1/2}.$$

Upper bounds for σ_k are given below.

Theorem 2. *For* $k = 2, \dots, n$

$$\sigma_k \le \frac{k-1}{k} n^2 \left[\sum_{m=1}^{k-1} \binom{n-k}{m-1} 2^{n-k+1-m} \sum_{2 \le i_1 < \ldots < i_m \le k} |\lambda_{i_1}| \cdots |\lambda_{i_m}| \right]^{1/(n-k+1)}.$$

In particular we have

$$\sigma_2 \le \frac{n^2}{2} 2^{(n-2)/(n-1)} |\lambda_2|^{1/(n-1)} = n^2 |\lambda_2/2|^{1/(n-1)},$$
$$\sigma_3 \le \frac{2}{3} n^2 \left[2^{n-3}(|\lambda_2| + |\lambda_3|) + (n-3) 2^{n-4} |\lambda_2||\lambda_3| \right]^{1/(n-2)}.$$

The results have both theoretical and applicational motivations. From theoretical point of view it is interesting to find a quantitative analogon of the Perron–Frobenius theorem, which can link eigenvalues of nonnegative matrices with uncoupling measures. One of the first result in this domain was the work by Fiedler [3], who bounded a subdominant eigenvalue of a doubly stochastic matrix in term of an irreducibility measure μ. In our work we generalize the Fiedler bounds for all stochastic matrices ($\sigma_2 = 2\mu$). As a byproduct we solve (Theorem 1) an open problem of Hartfiel and Meyer [4].

Recently Hartfiel proved an inequality (see Theorem 3 in [5]) which is equivalent to the following one

$$\nu_k \le 4^n n^{4n} |\lambda_k|^{1/n^{2n}}, \tag{1}$$

where

$$\nu_k := \min_{\substack{\emptyset \ne C_1, \dots, C_k \subset S \\ C_i \cap C_j = \emptyset \\ i,j = 1, \dots, k}} \max_{1 \le l \le k} \max_{i \in C_l} \sum_{j \notin C_l} p_{ij}.$$

It is easy to check that Theorem 2 is better then (1).

For numerical application it is interesting to bind an analysis of two different classes of algorithms for solving systems of linear equations related to Markov chains: iterative and aggregational one. Efficiency of the former class is characterised by eigenvalues of $\mathbf{P}$ while efficiency of the latter — in terms of uncoupling measures or in equivalent forms. The numerical experience points out that having subdominant eigenvalues of $\mathbf{P}$ near $\lambda = 1$ poses problems for numerical computation of characteristics of the Markov chain induced by an irreducible matrix $\mathbf{P}$. In particular the conditioning problem is large and iterative methods converge slowly (see [7] for more details). Most often in such cases the Markov chain is nearly completely decomposable and the aggregational algorithms produce small error. On the other hand if subdominant eigenvalues of $\mathbf{P}$ are separable from $\lambda = 1$ then iterative methods often converge rapidly but aggregational ones give large error. This paper is intended as a step towards theoretical explanation for this numerical experience.

It is worth to note that we prove the main results using the Matrix–Tree Theorem which represents principal minors of $\mathbf{I}-\mathbf{P}$ as sums of weights of directed forests (Lemma 1). This interesting method is probably unknown in the literature on numerical solutions of Markov chains. In the paper [8] we present other applications of the Matrix–Tree Theorem to computational problems related to Markov chains.

2. Proofs

Let $E \subseteq S \times S$. The (directed) *graph* with the *states* set S and the *edges* set E is, by definition, the pair $g := (S, E)$.

A pair $g_1 := (S_1, E_1)$ is called a *subgraph* of the graph g if $\emptyset \neq S_1 \subseteq S$ and $E_1 \subseteq E \cap (S_1 \times S_1)$.

A *path* from a state i to a state j in g_1 is, by definition, any finite sequence $i_0 = i, i_1, \ldots, i_k = j$ such that $i_m \in S_1$ and $(i_m, i_{m+1}) \in E$ for $m = 1, \ldots, k$. By a *cycle* in g_1 we mean a path from i to i.

A subgraph $f = (S_f, E_f)$ without cycles, in which from every state of S_f there is at most one outgoing edge is called the *forest*. The set of states $R \subseteq S_f$ from which there is no outgoing edge of the forest f is called the *root* of f.

It is easily seen that the root of f is nonempty and for every state $i \in S_f \backslash R$ there is only one path from i to R. We will denote by $F(R|A)$ the set of all forests with the root R and states set A. Set $F(R) := F(R|S)$.

Let $\mathbf{A} = (a_{ij})_{i,j \in S}$ be a square $n \times n$ matrix on $\mathbb{C}$.

The *weighted* graph induced by $\mathbf{A}$ is, by definition, a matrix $\mathbf{A}$ with the graph $g(\mathbf{A}) := (S, E)$, where $E = \{(i,j) \in S \times S : \ a_{ij} \neq 0\}$.

The (multiplicative) *weight* of a forest $f = (S_f, E_f)$ in $g(\mathbf{A})$ is defined to be

$$w(f) := \prod_{(i,j) \in E_f} (-a_{ij})$$

(set $w((S, \emptyset)) := 1$).

The weight of a set of forests F in $g(\mathbf{A})$ is defined to be

$$w(F) := \sum_{f \in F} w(f) \qquad (w(\emptyset) := 0).$$

If $F = F(R)$ for some $R \subset S$, we write $w(R)$ instead of $w(F(R))$, because the set R determines the set of all forests with R as the root.

The main tool we use is the following variant of the well–known Matrix Tree Theorem proved by Fiedler and Sedlacek [2] (see also [1] and the references given there).

Lemma 1 (Fiedler and Sedlacek 1958)**.** *Let* $\mathbf{L} := \mathbf{I} - \mathbf{P}$ *and* $R \subseteq S$. *Then*

$$\det \mathbf{L}(R) = w(R).$$

Let $\phi_{\mathbf{L}}(x) := \det(x\mathbf{I} - \mathbf{L}) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ be the characteristic polynomial of a matrix $\mathbf{L} := \mathbf{I} - \mathbf{P}$. It is clear that $a_0 = 0$ and $a_n = 1$.

Define

$$F^k := \bigcup_{\substack{R \subseteq S \\ |R| = k}} F(R) \qquad \text{for } k = 0, \ldots, n.$$

Lemma 2. *For* $k = 0, \ldots, n-1$

$$w(F^k) = \sum_{1 \le i_1 < \ldots < i_{n-k} \le n} \lambda_{i_1} \ldots \lambda_{i_{n-k}}.$$

Proof. By Lemma 1, formula 1.2.11 [6] and Viete's formulas, we obtain

$$w(F^k) = \sum_{\substack{R \subseteq S \\ |R| = k}} w(R) = \sum_{\substack{R \subseteq S \\ |R| = k}} \det \mathbf{L}(R) = (-1)^{s-k} a_k = \sum_{1 \le i_1 < \ldots < i_{n-k} \le n} \lambda_{i_1} \ldots \lambda_{i_{n-k}}.$$

□

Lemma 3. *For* $k = 2, \ldots, n$

$$w(F^k)\sigma_k \ge w(F^{k-1}).$$

Proof. Let $A_1, \ldots, A_k$ be nonempty proper subsets of S such that $A_i \cap A_j = \emptyset$ for $i \neq j$, $i, j \in S$ and

$$\sigma_k = \sum_{l=1}^{k} \sum_{\substack{i \in A_l \\ j \notin A_l}} p_{ij}. \tag{2}$$

We define a function $\varphi : F^{k-1} \to F^k$ erasing the edge of $f_1 \in F^{k-1}$ starting in one of the sets A_l $(l = 1, \ldots, k)$ which has the least starting point. Formally: for $f_1 \in F^{k-1}$ set

$$E^*_{f_1} := E_f \cap \bigcup_{l=1}^{k} (A_l \times \bar{A}_l),$$

$$i(f_1) := \min\{i \in S : \ (i,j) \in E^*_{f_1}\}.$$

By the definition of F^{k-1}every forest $f \in F^{k-1}$ has at least one edge starting from one of the sets A_l. Hence $E_{f_1} \neq \emptyset$. By the definition of a forest we know, that there are only one state $j \in S$ such that $(i(f_1), j) \in E_{f_1}$. Let us denote it by $j(f_1)$. The function φ is defined by formula

$$\varphi(f_1) = \big(S, E_{f_1} \setminus \{(i(f_1), j(f_1))\}\big).$$

Let us note that for $f_1 \in F^{k-1}$

$$w(f_1) = w(\varphi(f_1)) p_{i(f_1), j(f_1)}. \tag{3}$$

Since $w(F^{k-1}) = \sum_{f \in F^k} w(\varphi^{-1}(f))$ and $w(F^k) = \sum_{f \in F^k} w(f)$, to prove the lemma we need to show that for all $f \in F^k$

$$w(f)\sigma_k \geq w(\varphi^{-1}(f)).$$

Let $f \in F^k$ and $f_1, f_2 \in \varphi^{-1}(f)$. Obviously if $f_1 \neq f_2$ then $(i(f_1), j(f_1)) \neq (i(f_2), j(f_2))$.

Finally from (2) and (3) we have

$$\begin{aligned} w(f)\sigma_k = w(f) \left(\sum_{l=1}^{k} \sum_{\substack{i \in A_l \\ j \notin A_l}} p_{ij} \right) &\geq w(f) \left(\sum_{f_1 \in \varphi^{-1}(f)} p_{i(f_1), j(f_1)} \right) \\ = \sum_{f_1 \in \varphi^{-1}(f)} w(f_1) &= w(\varphi^{-1}(f)). \end{aligned}$$

□

Let us divide the set E into two parts: E_r "red edges" and E_b "black edges".

Lemma 4. *Let $R, A \subseteq S$ be nonempty sets such that $R \subseteq A$. We assume that for every nonempty $B \subseteq A \setminus R$ there is a black edge $(i, j) \in B \times (A \setminus B)$. Then there exists a forest $f \in F(R \big| A)$ which has only black edges.*

Proof. Let us define by induction:

$$\begin{aligned} &R_0 := R, \\ &R_k := \{i \in A : \ \exists j \in R_{k-1} \ (i, j) \in E_b\} \quad \text{for } k \geq 1. \end{aligned}$$

By assumption there is $k \in \mathbb{N}$ such that $A = R_0 \,\dot\cup \ldots \dot\cup\, R_k$ and $R_k \neq \emptyset$ for $k = 0, \ldots, K$. It follows that if we choose for every $i \in R_k$ $(k = 1, \ldots, K)$ one edge going from i to R_{k-1} we obtain a directed forest with the root R and the domain A. □

Lemma 5. *For $k = 2, \ldots, n$*

$$\mu_k \leq \left[w(F^{k-1})\right]^{1/(n-k+1)}.$$

Proof. Let $R := \{i_1, \ldots, i_{k-1}\}$ be a given subset of S,

$$\delta := w(F^{k-1}) \quad \text{and} \quad E_r := \{(i,j) : \; p_{ij} \leq \delta^{1/(n-k+1)}\}.$$

Since for every $f \in F(R)$ we have $w(f) \leq \delta$ and $|E_f| = n-k+1$, it follows that f has a red edge. By Lemma 4 there is a nonempty set $A_1 \subseteq S \setminus R$ satisfying the following condition:

$$\text{if} \quad (i,j) \in (A_1 \times (S \setminus A_1)) \cap E, \quad \text{then } (i,j) \in E_r. \tag{4}$$

We may assume that A_1 is a minimal set with property (4) in the order induced by inclusion. Therefore for every nonempty set $B \subset A_1$ there is a black edge outgoing from B to $A_1 \setminus B$. Choose $j_1 \in A_1$. By Lemma 4 it follows that there is a forest $f \in F(j_1 \big| A_1)$ which has only black edges. Thus we can find a forest $f \in F(j_1, i_2, \ldots, i_{k-1})$ such that $E_f = E_{f_1} \dot\cup E_{f_2}$, where $f_1 \in F(j_1 \big| A_1)$, f_1 has only black edges and $f_2 \in F(A_1^*)$, $A_1^* := A_1 \cup \{i_2, \ldots, i_{k-1}\}$. Hence every forest $f \in F(A_1^*)$ has a red edge. By Lemma 4 it follows that there exists nonempty set $A_2 \subseteq S \setminus A_1^*$ satisfying (4).

Again we may assume that A_2 is a minimal set with property (4) in the order induced by inclusion. Therefore for every nonempty set $B \subset A_2$ there is a black edge outgoing from B to $A_2 \setminus B$. Choose $j_2 \in A_2$. By Lemma 4 it follows that there is a forest $f \in F(j_2 \big| A_2)$ which has only black edges. We continue in this fashion obtaining nonempty and mutually disjoint sets $A_1, \ldots, A_{k-1}$ satisfying (4) which are minimal sets in the order induced by inclusion.

Now let $R := \{j_1, \ldots, j_{k-1}\}$ where $j_1 \in A_1, \ldots, j_{k-1} \in A_{k-1}$. By construction $E_f = E_{f_1} \dot\cup \ldots \dot\cup E_{f_{k-1}} \dot\cup E_{f_k}$, where $f_m \in F(j_m \big| A_m)$, f_m has only black edges ($m = 1, \ldots, k-1$) and $f_k \in F(A_1 \cup \ldots \cup A_{k-1})$. Then every forest $f \in F(A_1 \cup \ldots \cup A_{k-1})$ has a red edge. Hence, by Lemma 4 there is nonempty set $A_k \subseteq S \setminus (A_1 \cup \ldots \cup A_{k-1})$ satisfying (4).

Consequently, by definition of μ_k, we obtain

$$\mu_k \leq \max_{1 \leq l \leq k} \max_{\substack{i \in A_l \\ j \notin A_l}} p_{ij} \leq \delta^{1/(n-k+1)}.$$

□

Proof of Theorem 1. By Lemma 2 we have

$$\begin{aligned}\frac{w(F^k)}{w(F^1)} &= \sum_{2\le i_1<\ldots<i_{k-1}\le n} \frac{1}{\lambda_{i_1}\cdots\lambda_{i_{k-1}}} \\ &\le \sum_{2\le i_1<\ldots<i_{k-1}\le n} \frac{1}{|\lambda_{i_1}|\cdots|\lambda_{i_{k-1}}|} \\ &\le \binom{n-1}{k-1}\frac{1}{|\lambda_2|\cdots|\lambda_k|}.\end{aligned}$$

Applying Lemma 3 and Proposition 1 (i) we obtain

$$\begin{aligned}\binom{n-1}{k-1}^{-1}|\lambda_2|\cdots|\lambda_k| &\le \frac{w(F^1)}{w(F^k)} \\ &= \frac{w(F^{k-1})}{w(F^k)}\frac{w(F^{k-2})}{w(F^{k-1})}\cdots\frac{w(F^1)}{w(F^2)} \\ &\le \sigma_k\ldots\sigma_2.\end{aligned}$$

□

Proof of Theorem 2. By Lemma 2 we have that for $k = 1,\ldots,n-1$

$$w(F^{k-1}) \le \sum_{2\le i_1<\ldots<i_{n-k+1}\le n} |\lambda_{i_1}|\cdots|\lambda_{i_{n-k+1}}|.$$

Since $|\lambda_{k+1}|,\ldots,|\lambda_n| \le 2$, it follows that

$$w(F^{k-1}) \le \sum_{m=1}^{k-1}\binom{n-k}{m-1}2^{n-k+1-m}\sum_{2\le i_1<\ldots<i_m\le k}|\lambda_{i_1}|\cdots|\lambda_{i_m}|. \qquad (5)$$

From Proposition 2 and Lemma 5 we obtain

$$\sigma_k \le \frac{k-1}{k}n^2\left[w(F^{k-1})\right]^{1/(n-k+1)} \quad \text{for } k = 2,\ldots,n. \qquad (6)$$

Combining (5) with (6) we get the assertion. □

Acknowledgments. The author is greatly indebted to Prof. Lesław Gajek and Prof. Wojciech Niemiro for their helpful suggestions which make this work possible. Special thanks are due to Mrs Alicja Kossakowska–Kwiecień for preparing this paper.

References

[1] Chaiken, S., *A combinatorial proof of the all minors matrix tree theorem*, SIAM J. Alg. Disc. Meth. **3** (1982), 319–329.

[2] Fiedler, M. and Sedlácek, J., *O w–basich orientovaných grafu* (in Czech), Čas. Pest. Mat. **83** (1958), 214–225.

[3] Fiedler, M., *Bounds for eigenvalues of doubly stochastic matrices*, Linear Algebra Appl. **5** (1972), 299–310.

[4] Hartfiel, D.J. and Meyer, C.D., *On the structure of stochastic matrices with a subdominant eigenvalue near* 1, Linear Algebra Appl. **272** (1998), 193–203.

[5] Hartfiel, D.J., *Distribution of entries in a substochastic matrix having eigenvalues near* 1, (submitted for publication).

[6] Horn, R.A. and Johnson, Ch.R., *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.

[7] Meyer, C.D., *Sensitivity of the stationary distribution of a Markov chain*, SIAM J. Matrix Anal. Appl. **15**(3) 1994, 715–728.

[8] Pokarowski, P., *Directed forests with applications to algorithms related to Markov chains*, (to appear in Appl. Math.).

Piotr Pokarowski
Mathematical Institute
Polish Academy of Sciences
Śniadeckich 8
00-950 Warsaw, Poland
email: Pokar@impan.gov.pl