Linear Algebra and its Applications 431 (2009) 543–552

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Consistent behavior of certain perturbed determinants induced by graphs☆

Vladimir Ejov*, Giang T. Nguyen

Centre for Industrial and Applied Mathematics, University of South Australia, Australia

ARTICLE INFO

Article history:
Received 17 January 2008
Accepted 4 March 2009
Available online 21 April 2009

Submitted by H. Schneider

AMS classification:
60J10
05C45
11C20

Keywords:
Markov chains
Hamiltonian cycles
Deterministic probability transition matrices
Linear perturbations

ABSTRACT

We show that the determinant objective function introduced in Ejov et al. [V. Ejov, J. A. Filar, W. Murray, G.T. Nguyen, Determinants and longest cycles of graph, SIAM J. Discrete Math. 22 (33) (2008) 1215–1225] performs well under a certain symmetric linear perturbation. That means sub-graphs corresponding to Hamiltonian cycles of a given graph are maximizers over the hull of all sub-graphs with perturbation parameter $\varepsilon \in [0, 1)$. Note that in other optimization formulations (see, for example [V.S. Borkar, V. Ejov, J.A. Filar, Directed graphs, Hamiltonicity and doubly stochastic matrices, Random Structures Algorithms 25 (2004) 376–395; V. Ejov, J.A. Filar, M. Nguyen, Hamiltonian cycles and singularly perturbed Markov chains, Math. Oper. Res. 29 (1) (2004) 114–131; J.A. Filar, K. Liu, Hamiltonian cycle problem and singularly perturbed Markov decision process, in: Statistics, Probability and Game Theory: Papers in Honor of David Blackwell, IMS Lecture Notes – Monograph Series, USA, 1996]), ε in the corresponding perturbation was required to be significantly small.

1. Introduction

For more than two centuries, the Hamiltonian cycle problem has been attracting the attention and efforts of many researchers, who explored numerous approaches and employed different tools, mostly based on traditional graph theory and combinatorics, to truly understand the difficulty of this problem. In 1994, Filar and Krass [8] proposed a new line of research which approached the Hamiltonian cycle

☆ The authors gratefully acknowledge the Australian Research Council Discovery Grant Nos. DP0666632 and DP0984470.

* Corresponding author.
E-mail addresses: Vladimir.Ejov@unisa.edu.au (V. Ejov), Giang.Nguyen@unisa.edu.au (G.T. Nguyen).

0024-3795/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2009.03.005

problem from the perspective of Markov chains, thus converting the problem from whether a given graph possesses a Hamiltonian cycle (a cycle that goes through every vertex once) to minimizing an appropriate objective function over the set of all probability transition matrices of Markov chains feasible on the graph.

Since then, there has been a series of results on various objective functions and associated optimization models, whose global solutions help determine the Hamiltonicity of a given graph: top left element of fundamental matrices of Markov chains (see [9,6,3,1,2]), trace of fundamental matrices [7] and determinant of the inverse-like function of fundamental matrices [5]. While the majority of these papers focus on all feasible probability transition matrices under perturbations, [5] presents theoretical developments for unperturbed cases.

Consider a connected graph G of size N. Let $V(G)$ and $E(G)$ be the set of nodes and the set of arcs in G, respectively, where $N := |V(G)|$. Let $\mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of Markov chains entries of which represent the probabilities of traversing corresponding arcs of the graphs. Then, every entry in $\mathbf{P}$ is nonnegative. Furthermore, $\mathbf{P}$ is stochastic, that is, every row in $\mathbf{P}$ sums to 1. Formally,

$$\mathcal{F} := \left\{ \mathbf{P} | p_{ij} \geqslant 0; \sum_j p_{ij} = 1, \forall\ i;\ p_{ij} = 0, \forall\ (i,j) \notin E(G) \right\}.$$

Let $\mathcal{D} \subset \mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of deterministic Markov chains:

$$\mathcal{D} := \{ \mathbf{P} \subset \mathcal{F} | p_{ij} \in \{0,1\} \}.$$

A stochastic matrix $\mathbf{P}$ is said to be doubly stochastic if every column also adds up to 1. Let $\mathcal{DS} \subset \mathcal{F}$ be the set of all probability transition matrices $\mathbf{P}$ of doubly stochastic Markov chains:

$$\mathcal{DS} := \left\{ \mathbf{P} \subset \mathcal{F} | \sum_i p_{ij} = 1, \forall j \right\}.$$

For a perturbation parameter $\varepsilon \in [0,1)$, consider the symmetric linear perturbation:

$$\mathbf{P}^{\varepsilon} := (1-\varepsilon)\mathbf{P} + \frac{\varepsilon}{N}\mathbf{J}, \tag{1.1}$$

where $\mathbf{J}$ is an $N \times N$ matrix every entry of which is 1. Note that this symmetric linear perturbation preserves stochasticity and double stochasticity. Let $\mathbf{P}_{\mathrm{H}}^{\varepsilon}$ denote the perturbed probability transition matrix for any Hamiltonian cycle (HC).

The current paper is a continuation of [5]. We will show that the determinant objective function introduced in [5] performs well under the symmetric linear perturbation specified in (1.1). That means probability transition matrices corresponding to Hamiltonian cycles of a given graph are the maximizers over the set of all linearly perturbed feasible probability transition matrices. It is worthwhile to note that here, the perturbation parameter ε can take any value between 0 and 1. For many previous results obtained for other objective functions, ε was required to be significantly small. For example, the perturbation parameter was bounded from above by $1/N^2$ in [6], with N being the size of a given graph.

Section 2 briefly summarizes the results of [5] for unperturbed feasible probability transition matrices, whereas Sections 3 and 4 contain new theoretical derivations and construct an example for symmetric and asymmetric linear perturbations, respectively. Section 5 compares the robustness of the determinant function with that of other objective functions, under perturbations.

2. Determinants and longest cycles of graphs

For every stochastic $\mathbf{P}$, let $\mathbf{P}^*(\mathbf{P})$ be the stationary distribution matrix of a Markov chain represented by $\mathbf{P}$:

$$\mathbf{P}^*(\mathbf{P}) := \lim_{T\to\infty} \frac{1}{T+1} \sum_{i=1}^{T} \mathbf{P}^t,$$

where $\mathbf{P}^0 := \mathbf{I}$, and let $\mathbf{G}(\mathbf{P})$ be the fundamental matrix of the Markov chain:

$$\mathbf{G}(\mathbf{P}) := (\mathbf{I} - \mathbf{P} + \mathbf{P}^*(\mathbf{P}))^{-1}. \tag{2.1}$$

For doubly stochastic and irreducible probability transition matrix $\mathbf{P}$, matrix $\mathbf{P}^*(\mathbf{P})$ reduces to $\frac{1}{N}\mathbf{J}$. Consequently, the equality (2.1) becomes

$$\mathbf{G}(\mathbf{P}) := \left(\mathbf{I} - \mathbf{P} + \frac{1}{N}\mathbf{J}\right)^{-1}. \tag{2.2}$$

For $\alpha > 0$, introduce the matrix:

$$\mathbf{A}_\alpha(\mathbf{P}) := \mathbf{I} - \mathbf{P} + \alpha\mathbf{J}.$$

For convenience, we denote the probability transition matrix for any Hamiltonian cycle (HC) as $\mathbf{P}_{\mathrm{H}}$. Without loss of generality, assume that a HC is the standard cycle $1 \to 2 \to 3 \to \cdots \to N \to 1$.

Theorem 2.1 [5]. *For any graph G,*

$$\det_{\mathbf{P}\in\mathcal{F}} \mathbf{A}_\alpha(\mathbf{P}) \leqslant \alpha k N,$$

where k is the length of the longest cycle in G.

Corollary 2.2 [5]. *Let* $\alpha = \frac{1}{N}$, *then*

(i) *For any graph G,*

$$\det_{\mathbf{P}\in\mathcal{F}} \mathbf{A}(\mathbf{P}) \leqslant N.$$

(ii) *If G is a Hamiltonian graph, then equality is achieved and*

$$\det \mathbf{A}(\mathbf{P}) = N \Leftrightarrow \mathbf{P} = \mathbf{P}_{\mathrm{H}}.$$

(iii) *If G is a non-Hamiltonian graph, then*

$$\det_{\mathbf{P}\in\mathcal{F}} \mathbf{A}(\mathbf{P}) \leqslant N - 1.$$

These distinct upper bounds help us formulate the optimization problem:

$$\max_{\mathbf{P}\in\mathcal{F}} \mathbf{A}_\alpha(\mathbf{P}),$$

the global solutions of which determine the Hamiltonicity of a given graph G.

Remark 2.1. From this point onwards, for the sake of simplicity of the resulting formulae, we will focus on $\alpha = \frac{1}{N}$ exclusively. It can be shown, by analogous arguments, that the main results can be generalized.

3. Symmetric linear perturbation

Theorem 3.1. *Consider a Hamiltonian graph G. For any probability transition matrix* $\mathbf{P} \in \mathcal{F}$,

$$\max_{\mathbf{P}^\varepsilon\in\mathcal{F}} \det \mathbf{A}(\mathbf{P}^\varepsilon) = \mathbf{A}\left(\mathbf{P}^\varepsilon_{\mathrm{H}}\right),$$

for some $\mathbf{P}^\varepsilon_{\mathrm{H}}$ *corresponding to a Hamiltonian cycle.*

The proof of Theorem 3.1 is based on four lemmata. These lemmata are organized as follows: Lemma 3.2 shows that for any $\mathbf{P} \in \mathcal{F}$, the determinant of $\mathbf{A}(\mathbf{P}^\varepsilon)$ is bounded above by the corresponding

determinant of some perturbed deterministic **P**. Lemmata 3.3,3.4,3.5 provide the exact formulae for three disjoint exhaustive classes of deterministic probability transition matrices: Hamiltonian cycles, probability transition matrices that contain one single cycle of length $k < N$, and probability transition matrices that contain two or more disjoint cycles, respectively. Theorem 3.1 then follows from the comparison of these three expressions.

Lemma 3.2. *For any probability transition matrix* $\mathbf{P} \in \mathcal{F}$, *there exists a deterministic probability transition matrix* $\mathbf{D} \in \mathcal{D}$ *such that*

$$\det \mathbf{A}(\mathbf{D}^\varepsilon) \geqslant \det \mathbf{A}(\mathbf{P}^\varepsilon).$$

Proof. Although the proof is similar to the proof of Proposition 4 in [5], we reproduce it for the sake of completeness. Since $\mathcal{F}$ is compact, $\max_{\mathcal{F}} \det \mathbf{A}(\mathbf{P}^\varepsilon)$ exists, let

$$\widehat{\mathcal{P}}^\varepsilon = \left\{ \arg \max_{\mathbf{P}^\varepsilon \in \mathcal{F}} \det \mathbf{A}(\mathbf{P}^\varepsilon) \right\}$$

and $\widehat{\mathcal{P}}$ is the corresponding set of unperturbed probability transition matrices, that is, $\widehat{\mathcal{P}} := \widehat{\mathcal{P}}^0$. Consider $\widehat{\mathbf{P}} \in \widehat{\mathcal{P}}$ with the minimal total number of randomizations in its rows compared to the other probability transition matrices in $\widehat{\mathcal{P}}$, and the corresponding perturbed probability transition matrix $\widehat{\mathbf{P}}^\varepsilon \in \widehat{\mathcal{P}}^\varepsilon$. If $\widehat{\mathbf{P}}$ is a deterministic probability transition matrix $\mathbf{D} \in \mathcal{D}$, there is nothing to prove.

Suppose $\widehat{\mathbf{P}} \notin \mathcal{D}$, then $\widehat{\mathbf{P}}$ contains a row, say row 1, where a randomization occurs. Let this row be

$$\begin{bmatrix} \cdots & a & \cdots & b & \cdots & c & \cdots \end{bmatrix}, \quad a, b > 0.$$

Consider $\widehat{\mathbf{D}}_\nu \in \mathcal{F}$ that coincides with $\widehat{\mathbf{P}}$ in rows $2, \ldots, N$ and whose first row is

$$\begin{bmatrix} \cdots & \nu & \cdots & \frac{(1-\nu)b}{1-a} & \cdots & \frac{(1-\nu)c}{1-a} & \cdots \end{bmatrix}, \quad 0 \leqslant \nu \leqslant 1$$

and the corresponding perturbed probability transition matrix $\widehat{\mathbf{D}}^\varepsilon_\nu$. Note that $\widehat{\mathbf{D}}_{\nu=a} = \widehat{\mathbf{P}}$. As $\det \mathbf{A}(\widehat{\mathbf{D}}^\varepsilon_\nu)$ is a linear function in ν, $\max_{\nu\in[0,1]} \det \mathbf{A}(\widehat{\mathbf{D}}^\varepsilon_\nu)$ occurs either at $\nu = 0$ or $\nu = 1$, or else $\det \mathbf{A}(\widehat{\mathbf{D}}^\varepsilon_\nu) = \det \mathbf{A}(\widehat{\mathbf{P}}^\varepsilon)$ for all $\nu \in [0, 1]$.

In each of these situations, $\widehat{\mathbf{D}} = \widehat{\mathbf{D}}_{\nu=0}$ or $\widehat{\mathbf{D}}_{\nu=1}$ has at least one more zero than $\widehat{\mathbf{P}}$, and $\det \mathbf{A}(\widehat{\mathbf{D}}^\varepsilon) \geqslant \det \mathbf{A}(\widehat{\mathbf{P}}^\varepsilon)$ so $\widehat{\mathbf{D}}^\varepsilon \in \{\arg\max_{\mathbf{P}^\varepsilon\in\mathcal{F}} \det \mathbf{A}(\mathbf{P}^\varepsilon)\}$. This leads to a contradiction to the assumption that $\widehat{\mathbf{P}}$ has the minimal total number of randomizations in its rows compared to the other probability transition matrices in $\widehat{\mathcal{P}}$. Consequently, $\widehat{\mathbf{P}}$ is deterministic, and this concludes the proof. □

Lemma 3.3. *Consider a Hamiltonian graph G. For any Hamiltonian cycle* $\mathbf{P}_\mathrm{H} \in \mathcal{D}$,

$$\det \mathbf{A}(\mathbf{P}^\varepsilon_\mathrm{H}) = \frac{1 - (1-\varepsilon)^N}{\varepsilon}. \tag{3.1}$$

Proof. For $i = 1, \ldots, N$, let λ_i be eigenvalues of $\mathbf{P}_\mathrm{H}$. By the proof of Lemma 4.2 in [5], every $\lambda_i \neq 1$ is also an eigenvalue of $\mathbf{P}_\mathrm{H} - \frac{1}{N}\mathbf{J}$ and vice versa. For every $\mathbf{P}_\mathrm{H}$, eigenvalues λ_i are Nth roots of unity (see, for example, [4]). Therefore, there exists only one eigenvalue of 1 for $\mathbf{P}_\mathrm{H}$, which we denote λ_N. It is straightforward to see that eigenvector $\mathbf{e} = (1, \ldots, 1)^\mathrm{T}$ corresponds to this eigenvalue, as $\mathbf{Pe} = \mathbf{e}$. This eigenvector $\mathbf{e}$ is also an eigenvector of $\mathbf{P}_\mathrm{H} - \frac{1}{N}\mathbf{J}$, associated with eigenvalue 0, as $\left(\mathbf{P}_\mathrm{H} - \frac{1}{N}\mathbf{J}\right)\mathbf{e} = \mathbf{0}$. Hence,

$$\begin{aligned} \det \mathbf{A}(\mathbf{P}^\varepsilon_\mathrm{H}) &= \det\left[\mathbf{I} - \left((1-\varepsilon)\mathbf{P}_\mathrm{H} + \frac{\varepsilon}{N}\mathbf{J}\right) + \frac{1}{N}\mathbf{J}\right] \\ &= \det\left[\mathbf{I} - (1-\varepsilon)\left(\mathbf{P}_\mathrm{H} - \frac{1}{N}\mathbf{J}\right)\right] = \prod_{i=1}^{N-1}(1 - (1-\varepsilon)\lambda_i) \end{aligned}$$

$$= 1 - (1-\varepsilon)\sum_{i=1}^{N-1}\lambda_i + (1-\varepsilon)^2\sum_{\substack{i\geqslant j\\ ij=1}}^{N-1}\lambda_i\lambda_j + \cdots + (-1)^{N-1}(1-\varepsilon)^{N-1}\prod_{i=1}^{N-1}\lambda_i$$

$$= 1 - (1-\varepsilon)q_1(\lambda_i) + (1-\varepsilon)^2 q_2(\lambda_i) + \cdots + (-1)^{N-1}(1-\varepsilon)^{N-1}q_{N-1}(\lambda_i), \tag{3.2}$$

where q_i is the ith elementary symmetric polynomial in $\lambda_i, i = 1,\ldots,N-1$.

From the proof of Proposition 1 in [5], $q_1 = -1, q_2 = 1,\ldots,q_{N-1} = (-1)^{N-1}$, so (3.2) becomes

$$\det \mathbf{A}(\mathbf{P}_{\mathrm{H}}^{\varepsilon}) = 1 + (1-\varepsilon) + (1-\varepsilon)^2 + \cdots + (1-\varepsilon)^{N-1} = \frac{1-(1-\varepsilon)^N}{1-(1-\varepsilon)} = \frac{1-(1-\varepsilon)^N}{\varepsilon},$$

which asymptotically tends to N as $\varepsilon \to 0$. □

Lemma 3.4. *For any $\mathbf{P} \in \mathcal{D}$ that contains exactly one cycle of length $2 \leqslant k < N$,*

$$\det \mathbf{A}(\mathbf{P}^{\varepsilon}) = \frac{1-(1-\varepsilon)^k}{\varepsilon}. \tag{3.3}$$

Proof. As in Lemma 3.3, we have

$$\mathbf{A}(\mathbf{P}^{\varepsilon}) = \mathbf{I} - \left[(1-\varepsilon)\mathbf{P} + \frac{\varepsilon}{N}\mathbf{J}\right] + \frac{1}{N}\mathbf{J} = \mathbf{I} + (1-\varepsilon)\left[-\mathbf{P} + \frac{1}{N}\mathbf{J}\right].$$

Let λ_i be the eigenvalues of $-\mathbf{P} + \frac{1}{N}\mathbf{J}$. From the proof of Lemma 1 in [5], the last $N-k+1$ eigenvalues $\lambda_k = \lambda_{k+1} = \ldots = \lambda_N = 0$, and the first $k-1$ eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ are identical to those of $\mathbf{P}_{\mathrm{H}_k}$, where $\mathbf{P}_{\mathrm{H}_k}$ is an k by k matrix containing a single cycle of length k.

Let μ_i be the eigenvalues of $(1-\varepsilon)\left[-\mathbf{P} + \frac{1}{N}\mathbf{J}\right]$, then

$$\det \mathbf{A}(\mathbf{P}^{\varepsilon}) = \prod_{i=1}^{N}(1-\mu_i) = \prod_{i=1}^{N}(1-(1-\varepsilon)\lambda_i) = \prod_{i=1}^{k-1}(1-(1-\varepsilon)\lambda_i) = \frac{1-(1-\varepsilon)^k}{\varepsilon},$$

where the last equality follows from Lemma 3.3, for $N = k$. The above tends to k as $\varepsilon \to 0$. □

Lemma 3.5. *For any $\mathbf{P} \in \mathcal{D}$ that contains $m \geqslant 2$ disjoint cycles of lengths $k_i \geqslant 2, \sum_i k_i \leqslant N$,*

(i) *the multiplicity of eigenvalue 1 for $\mathbf{P}$ is m. The structure of eigenvectors corresponding to these eigenvalues are given in the proof that follows, and*
(ii) *the following equality holds*

$$\det \mathbf{A}(\mathbf{P}^{\varepsilon}) = \varepsilon^{m-1}\prod_{i=1}^{m}\frac{1-(1-\varepsilon)^{k_i}}{\varepsilon}, \quad i = 1,\ldots,m. \tag{3.4}$$

Proof. (i) Consider $\mathbf{P} \in \mathcal{D}$ that contains $m \geqslant 2$ disjoint cycles of length $k_i \geqslant 2, \sum_i k_i \leqslant N$, and the cycles are $1 \to 2 \to \cdots \to k_1 \to 1, (k_1+1) \to (k_1+2) \to \cdots \to (k_1+k_2) \to (k_1+1), \ldots,$ $\left(\left(\sum_{j=1}^{k_{m-1}} k_j\right)+1\right) \to \left(\sum_{j=1}^{k_{m-1}} k_j + 2\right) \to \cdots \to \left(\sum_{j=1}^{k_m} k_j\right) \to \left(\left(\sum_{j=1}^{k_{m-1}} k_j\right)+1\right)$.

Moreover, for the $N - \sum_i k_i$ vertices that are not on these m disjoint cycles, assume that t_j of them are connected to the jth cycle, either by one or more edges, for $j = 1,\ldots,m$, and $\sum_j t_j = N - \sum_i k_i$. Without loss of generality, assume that $\mathbf{P}$ has the following structure, where vertices that are not on any cycle are labeled to be from $\sum_i k_i + 1$ to N,

$$\left[\begin{array}{c|c|c|c|c}
\begin{matrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{matrix} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline
\mathbf{0} & \begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{matrix} & \cdots & \mathbf{0} & \mathbf{0} \\ \hline
\cdots & \cdots & \cdots & \cdots & \cdots \\ \hline
\begin{matrix} & & & \\ & & & \\ & & & \\ & & 1 & \end{matrix} & \mathbf{0} & \cdots & \begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{matrix} & \mathbf{0} \\ \hline
\mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \begin{matrix} 0 & 1 & \\ & \ddots & 1 \\ 0 & & 0 \end{matrix}
\end{array}\right]. \tag{3.5}$$

For the ith cycle, $i = 1, \ldots, m$, we construct a vector $\underline{w}_i$ as follows:

- every entry of $\underline{w}_i$ corresponding to a vertex on this cycle is 1,
- every entry corresponding to a vertex not on the cycle that is connected, by one or more edges, to this cycle is also 1, and
- every other entry is 0.

It is easy to verify that $\mathbf{P}\underline{w}_i = \underline{w}_i$, for $i = 1, \ldots, m$. Consequently, $\underline{w}_i$ are eigenvectors of $\mathbf{P}$ corresponding to the eigenvalue 1. Moreover, it is clear that these eigenvectors are linearly independent and the multiplicity of the eigenvalue 1 of $\mathbf{P}$ is m.

Let n_i be the number of units in $\underline{w}_i$, $i = 1, \ldots, m$. For the purpose of demonstrating the relationship between eigenvalues of $\mathbf{P}$ and of $-\mathbf{P} + \frac{1}{N}\mathbf{J}$ later on, we decompose $\mathbb{R}^N$ as $\mathrm{span}\{\underline{e}\} \bigoplus \ker(\mathbf{J})$. Consequently, instead of vectors $\underline{w}_i$, $i = 1, \ldots, m$, it is convenient to consider their linear combinations as follows: $\underline{u}_m := \sum_{i=1}^{m} \underline{w}_i = \underline{e}$, and $\underline{u}_i := \underline{w}_m - \frac{n_m}{n_i}\underline{w}_i$, for $i = 1, \ldots, m-1$.

(ii) Without loss of generality, assume that $\mathbf{P} \in \mathcal{D}$ contains only two cycles of lengths $2 \leqslant k_1, k_2 < N$ respectively, $k_1 + k_2 < N$, and $\mathbf{P}$ is of the generic structure:

$$\left[\begin{array}{c|c|c}
\begin{matrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{matrix} & \mathbf{0} & \mathbf{0} \\ \hline
\mathbf{0} & \begin{matrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{matrix} & \mathbf{0} \\ \hline
\begin{matrix} & & & \\ & & & \\ & & & \\ & & 1 & \end{matrix} & \mathbf{0} & \begin{matrix} 0 & 1 & \\ & \ddots & 1 \\ 0 & & 0 \end{matrix}
\end{array}\right]. \tag{3.6}$$

Firstly, let $\mathbf{E} = \{\underline{e}_i\}$ be the standard basis of $\mathbb{R}^N$ where for each $\underline{e}_i$, the ith entry is unity and all other entries are zeroes. With respect to the decomposition of $\mathbb{R}^N = \text{span}\{\underline{e}\} \bigoplus \ker(\mathbf{J})$, we consider a new basis $\mathbf{V}$ of $\mathbb{R}^N$ that consists of

$$\begin{aligned}
&\underline{v}_1 = \underline{e}_2 - \underline{e}_1, \ldots, \underline{v}_{k_1-1} = \underline{e}_{k_1} - \underline{e}_1;\\
&\underline{v}_{k_1} := \underline{u}_1;\\
&\underline{v}_{k_1+1} := \underline{e}_{k_1+2} - \underline{e}_{k_1+1}, \ldots, \underline{v}_{k_1+k_2-1} := \underline{e}_{k_1+k_2+1} - \underline{e}_{k_1+1};\\
&\underline{v}_{k_1+k_2} := \underline{u}_2 = \underline{e}; \quad \underline{v}_{k_1+k_2+1} := \underline{e}_{k_1+k_2+1}, \ldots, \underline{v}_N := \underline{e}_N.
\end{aligned}$$

Then, $\underline{v}_{k_1+k_2} = \underline{e} \in \ker\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$. Also,

$$\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)\underline{v}_{k_1+k_2+1} = -\mathbf{P}\underline{e}_{k_1+k_2+1} + \frac{1}{N}\mathbf{J}\underline{e}_{k_1+k_2+1} = \frac{1}{N}\underline{e}.$$

Similarly, for $k_1 + k_2 + 2 \leqslant i \leqslant N$,

$$\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)\underline{v}_i = \gamma_i \underline{v}_{i-1} + \frac{1}{N}\underline{e}, \quad \text{for } \gamma_i \in \{-1, 0\}$$

and consequently,

$$\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)^{N-(k_1+k_2)+1} \underline{v}_i = \underline{0}.$$

Therefore, vectors $\underline{v}_{k_1+k_2}, \ldots, \underline{v}_N$ form a basis of $\ker\left(\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)^{N-(k_1+k_2)+1}\right)$. Hence, 0 is an eigenvalue of $\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$, and consequently of $(1-\varepsilon)\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$, of multiplicity at least $N - (k_1 + k_2) + 1$. This means that 1 is an eigenvalue of $\mathbf{A}(\mathbf{P})$, and consequently of $\mathbf{A}(\mathbf{P}^\varepsilon)$, of multiplicity at least $N - (k_1 + k_2) + 1$.

It is straightforward to check that vector $\underline{v}_{k_1}$ is an eigenvector of $\mathbf{A}(\mathbf{P}^\varepsilon)$ corresponding to an eigenvalue of ε. This is because $\mathbf{I}\underline{v}_{k_1} = \underline{v}_{k_1}, \mathbf{J}\underline{v}_{k_1} = \underline{0}$ and $\mathbf{P}^\varepsilon \underline{v}_{k_1} = [(1-\varepsilon)\mathbf{P} + \frac{\varepsilon}{N}\mathbf{J}]\underline{v}_{k_1} = (1-\varepsilon)\underline{v}_{k_1}$. Therefore,

$$\mathbf{A}(\mathbf{P}^\varepsilon)\underline{v}_{k_1} = \left[\mathbf{I} - \mathbf{P}^\varepsilon + \frac{1}{N}\mathbf{J}\right]\underline{v}_{k_1} = \varepsilon \underline{v}_{k_1}.$$

Let μ_i be the eigenvalues of $\mathbf{A}(\mathbf{P}^\varepsilon)$, and λ_i be the eigenvalues of $\mathbf{P}$, and consequently $(1-\varepsilon)\lambda_i$ be the eigenvalues of $\mathbf{P}^\varepsilon$. So far, we have shown that $\mu_i = 1$, for $i = k_1 + k_2, \ldots, N$, and $\mu_i = \varepsilon$ for $i = k_1$.

It remains to show that for $i = 1, \ldots, k_1 - 1$, λ_i coincide with the eigenvalues of $\mathbf{P}_{\mathrm{H}_{k_1}}$ (with eigenvalue of 1 excluded), and for $i = k_1 + 1, \ldots, k_1 + k_2 - 1$, λ_i coincide with the eigenvalues of $\mathbf{P}_{\mathrm{H}_{k_2}}$ (also with eigenvalue of 1 excluded), where $\mathbf{P}_{\mathrm{H}_{k_1}}$ and $\mathbf{P}_{\mathrm{H}_{k_2}}$ are probability transition matrices corresponding to Hamiltonian cycles in graphs of sizes k_1 and k_2, respectively.

Indeed, $\mathbf{J}\underline{v}_i = 0$, for $i = 1, \ldots, k_1 - 1$. Hence, according to the matrix structure illustrated in (3.6),

$$\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)\underline{v}_1 = -\mathbf{P}(\underline{e}_2 - \underline{e}_1) = -\underline{e}_1 + \underline{e}_{k_1} + \underline{v}_1^*, \quad \text{where } \underline{v}_1^* \in \text{span}\{v_{k_1+k_2}, \ldots, v_N\}.$$

Analogously,

$$\begin{aligned}
\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)\underline{v}_i &= -\mathbf{P}(\underline{e}_{i+1} - \underline{e}_1) = -\underline{e}_i + \underline{e}_{k_1} + \underline{v}_i^*,\\
&\text{where } \underline{v}_i^* \in \text{span}\{\underline{v}_{k_1+k_2}, \ldots, \underline{v}_N\}, \quad i = 2, \ldots, k_1 - 1.
\end{aligned}$$

Therefore, we observe that the $(k_1 - 1) \times (k_1 - 1)$ principal leading minor for $\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$ is the same as the $(k_1 - 1) \times (k_1 - 1)$ principal leading minor of $-\mathbf{P}_{\mathrm{H}_{k_1}} + \frac{1}{k_1}\mathbf{J}_{k_1}$, if we choose bases $\mathbf{V}$ and

$\{\underline{e}_2^{(k_1)} - \underline{e}_1^{(k_1)}, \ldots, \underline{e}_{k_1-1}^{(k_1)} - \underline{e}_1^{(k_1)}, \underline{e}^{(k_1)}\}$ respectively, where $\mathbf{J}_{k_1}$ is the $k_1 \times k_1$ matrix of units; $\underline{e}_j^{(k_1)}$ is the jth vector of the standard basis of $\mathbb{R}^{k_1}$ and $\underline{e}^{(k_1)} = [1, 1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^{k_1}$.

Thus, in the basis $\mathbf{V}$, matrix $\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$ has the form:

$$\left[\begin{array}{c|c} -\mathbf{P}_{\mathrm{H}_{k_1}} + \frac{1}{k_1}\mathbf{J}_{k_1} & \mathbf{0} \\ \hline *_1 & *_2 \end{array}\right],$$

where $*_1$ denotes some $(N - k_1) \times k_1$ matrix and $*_2$ denotes some $(N - k_1) \times (N - k_1)$ matrix. As it can be easily verified that $-\mathbf{P}_{\mathrm{H}_{k_1}} + \frac{1}{k_1}\mathbf{J}_{k_1}$ is diagonalizable, the eigenvalues λ_i of $\mathbf{P}$, for $i = 1, \ldots, k_1 - 1$ (which are also eigenvalues of $\left(-\mathbf{P} + \frac{1}{N}\mathbf{J}\right)$) coincide with the eigenvalues of $\mathbf{P}_{\mathrm{H}_{k_1}}$ (with eigenvalue of 1 excluded).

Analogously, we can show that for $i = k_1 + 1, \ldots, k_1 + k_2 - 1$, the eigenvalues λ_i of $\mathbf{P}$ coincide with the eigenvalues of $\mathbf{P}_{\mathrm{H}_{k_2}}$ (with eigenvalue of 1 excluded), by swapping the two diagonal blocks of $\mathbf{P}$ corresponding to the two cycles.
Therefore,

$$\det \mathbf{A}(\mathbf{P}^{\varepsilon}) = \varepsilon \prod_{i=1}^{k_1-1} (1 - \lambda_i) \prod_{i=k_1+1}^{k_1+k_2-1} (1 - \lambda_i) = \varepsilon \left[\frac{1 - (1-\varepsilon)^{k_1}}{\varepsilon}\right] \left[\frac{1 - (1-\varepsilon)^{k_2}}{\varepsilon}\right],$$

where the last equality follows from Lemma 3.3, for $N = k_1$ and $N = k_2$. This asymptotically tends to 0 as $\varepsilon \to 0$. We remark that the above formula can be easily generalized for the arbitrary number of disjoint cycles. □

Proof of Theorem 3.1. We need to show that for any stochastic $\mathbf{P} \in \mathcal{F}$,

$$\arg \max_{\mathbf{P}^{\varepsilon} \in \mathcal{F}} \det \mathbf{A}(\mathbf{P}^{\varepsilon}) = \mathbf{P}_{\mathrm{H}}^{\varepsilon}.$$

In order to prove this part, all we need to show is, for $i = 1, \ldots, m$; $k, \ell_i < N$ and $\sum_i \ell_i \leqslant N$,

$$\frac{1 - (1-\varepsilon)^N}{\varepsilon} > \frac{1 - (1-\varepsilon)^k}{\varepsilon} \quad \text{and} \quad \frac{1 - (1-\varepsilon)^N}{\varepsilon} > \varepsilon^{m-1} \prod_{i=1}^{m} \frac{1 - (1-\varepsilon)^{\ell_i}}{\varepsilon}.$$

The first inequality is straightforward, as $(1-\varepsilon)^N < (1-\varepsilon)^k$, for $k < N$. For the second inequality, it is sufficient to prove that

$$1 - (1-\varepsilon)^N > \prod_{i=1}^{m} [1 - (1-\varepsilon)^{\ell_i}].$$

This follows as

$$\prod_{i=1}^{m} [1 - (1-\varepsilon)^{\ell_i}] \leqslant 1 - (1-\varepsilon)^{\ell_1} < 1 - (1-\varepsilon)^N.$$

This concludes the proof of Theorem 3.1. □

4. Asymmetric linear perturbation

Following [8], for $\varepsilon \in (0, 1)$, consider the asymmetric linear perturbation:

$$\mathbf{P}^{\varepsilon} := (1-\varepsilon)\mathbf{P} + \varepsilon \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \mathbf{P} + \varepsilon \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}. \tag{4.1}$$

We will include an example to demonstrate that the results of Theorem 3.1 do not hold for the aforementioned asymmetric linear perturbation.

Example 4.1. Let $\mathbf{P}_{\mathrm{H}}$ be the probability transition matrix for a standard HC of length $N = 20$:

$$\mathbf{P}_{\mathrm{H}} = \begin{bmatrix} 0 & 1 & & & & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \ddots & 1 \\ 1 & & & & & 0 \end{bmatrix}.$$

Let $\mathbf{P}_1$ be the probability transition matrix that traces out a single cycle of length $k = 19$ corresponding to a standard HC of length 19 and the last vertex connects to the first vertex of this HC:

$$\mathbf{P}_1 = \left[\begin{array}{ccccc|c} 0 & 1 & & & & 0 \\ & \ddots & \ddots & & & \vdots \\ & & \ddots & & & \vdots \\ & & & \ddots & 1 & \vdots \\ 1 & & & & 0 & 0 \\ \hline 1 & & & & 0 & 0 \end{array}\right].$$

Applying the asymmetric linear perturbation in (4.1), computation by MAPLE shows that for $\varepsilon = 0.8$, $\det \mathbf{A}(\mathbf{P}_{\mathrm{H}}^{\varepsilon}) \approx 2.249995$ and $\det \mathbf{A}(\mathbf{P}_1^{\varepsilon}) \approx 2.249999$.

We speculate one possible reason for Theorem 3.1 to fail for the asymmetric linear perturbation specified in (4.1) for large values of ε. It might be because this asymmetric linear perturbation is, by design, biased towards vertex 1 in any given graph. Therefore, for large values of ε, the uniformly distributed probability transition matrices corresponding to Hamiltonian cycles, once asymmetrically perturbed, lose their optimal property for the determinant function, which is a symmetric objective function. This is not the case for the symmetric linear perturbation.

5. Behaviors of different objective functions under perturbation

Consider two other objective functions $g_{11}(\mathbf{P}) = [\mathbf{G}(\mathbf{P})]_{11}$ and $\mathrm{Tr}[\mathbf{G}(\mathbf{P})]$ introduced in [9,7], respectively.

Recall that from (2.1), the fundamental matrix $\mathbf{G}(\mathbf{P})$ is defined as

$$\mathbf{G}(\mathbf{P}) := (\mathbf{I} - \mathbf{P} + \mathbf{P}^*(\mathbf{P}))^{-1}.$$

Then, the first objective function is defined as the top left element of the fundamental matrix $\mathbf{G}(\mathbf{P})$: $g_{11}(\mathbf{P}) = [\mathbf{G}(\mathbf{P})]_{11}$. The second objective function is defined as the trace of the fundamental matrix $\mathbf{G}(\mathbf{P})$:

$$\mathrm{Tr}[\mathbf{G}(\mathbf{P})] = \sum_{i=1}^{N} [\mathbf{G}(\mathbf{P})]_{ii}. \tag{5.1}$$

From now on, by symmetric linear perturbation, we are referring to the symmetric linear perturbation defined in (1.1), and by asymmetric linear perturbation, we are referring to the asymmetric linear perturbation defined in (4.1).

In [6], it was proven that the minimizers of $g_{11}(\mathbf{P}^{\varepsilon})$ over the set of all perturbed deterministic probability transition matrices $\mathcal{D}$ correspond to Hamiltonian cycles, for $\varepsilon \in [0, 1/N^2)$ and with the asymmetric linear perturbation. In [1], the result was generalized to the set of all doubly stochastic matrices, for sufficiently small ε and with the symmetric linear perturbation. In [7], the following

example was constructed to demonstrate that this result does not hold over the set of stochastic matrices, for a large value of ε and with both symmetric and asymmetric linear perturbations.

Example 5.1. Consider the following probability transition matrices

$$\mathbf{P}_1 = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{P}_{\mathrm{H}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Applying the symmetric linear perturbation, computation by MAPLE shows that for $\varepsilon = 0.8$, $g_{11}(\mathbf{P}_1^{\varepsilon}) = 0.950$ and $g_{11}(\mathbf{P}_{\mathrm{H}}^{\varepsilon}) = 0.917$. Applying the asymmetric linear perturbation, computation by MAPLE shows that for $\varepsilon = 0.8$, $g_{11}(\mathbf{P}_1^{\varepsilon}) = 0.750$ and $g_{11}(\mathbf{P}_{\mathrm{H}}^{\varepsilon}) = 0.752$.

On the other hand, in the same paper [7], it was proven that the minimizers of $\mathrm{Tr}[\mathbf{G}(\mathbf{P}^{\varepsilon})]$ over the set of all doubly stochastic matrices correspond to Hamiltonian cycles, for $\varepsilon \in [0, 1)$ and with the symmetric linear perturbation. It was also shown that this result holds over the set of all stochastic matrices for $\varepsilon = 0$, that is, without any perturbation.

The paper includes a conjecture that with the symmetric linear perturbation, the result still holds for $\varepsilon \in [0, 1)$. Note that for all the examples that we considered, the inequality $\mathrm{Tr}[\mathbf{G}(\mathbf{P}_{\mathrm{H}}^{\varepsilon})] \leqslant \mathrm{Tr}[\mathbf{G}(\mathbf{P}^{\varepsilon})]$ holds, where $\mathbf{P}$ does not correspond to a Hamiltonian cycle, for both symmetric and asymmetric linear perturbations. However, the result analogous to Theorem 3.1 for the objective function $\mathrm{Tr}[\mathbf{G}(\mathbf{P}^{\varepsilon})]$ remains an intriguing open problem.

Acknowledgements

The authors would like to thank J. Filar and M. Haythorpe for numerous useful discussions. We are also grateful to an anonymous referee whose comments significantly improved the manuscript.

References

[1] V.S. Borkar, V. Ejov, J.A. Filar, Directed graphs, Hamiltonicity and doubly stochastic matrices, Random Structures Algorithms, 25 (2004) 376–395.
[2] V.S. Borkar, V. Ejov, J.A. Filar, On the Hamiltonicity gap and doubly stochastic matrices, Random Structures Algorithms, doi:10.1002/rsa.20237.
[3] V. Ejov, J.A. Filar, J. Gondzio, An interior point heuristic for the Hamiltonian cycle problem via Markov decision processes, J. Global Optim. 29 (2004) 315–334.
[4] V.Ejov, J.A. Filar, S.K. Lucas, J.L. Nelson, Solving the Hamiltonian cycle problem using symbolic determinants, Taiwanese J. Math. 10 (2006) 327–338.
[5] V. Ejov, J.A. Filar, W. Murray, G.T. Nguyen, Determinants and longest cycles of graphs, SIAM J. Discrete Math. 22 (33) (2008) 1215–1225.
[6] V. Ejov, J.A. Filar, M. Nguyen, Hamiltonian cycles and singularly perturbed Markov chains, Math. Oper. Res. 29 (1) (2004) 114–131.
[7] N. Litvak, V. Ejov, Markov chains and optimality of the Hamiltonian cycle, Math. Oper. Res. 3 (1) (2009) 71–82.
[8] J.A. Filar, D. Krass, Hamiltonian cycles and Markov chains, Math. Oper. Res. 19 (1994) 223–237.
[9] J.A. Filar, K. Liu, Hamiltonian cycle problem and singularly perturbed Markov decision process, in: Statistics, Probability and Game Theory: Papers in Honor of David Blackwell, IMS Lecture Notes – Monograph Series, USA, 1996.