# On the Determinant and its Derivatives of the Rank-one Corrected Generator of a Markov Chain on a Graph.

**J.A. Filar, M. Haythorpe, W. Murray**

**Abstract** We present an algorithm to find the determinant and its first and second derivatives of a rank-one corrected generator matrix of a doubly stochastic Markov chain. The motivation arises from the fact that the global minimiser of this determinant solves the Hamiltonian cycle problem. It is essential for algorithms that find global minimisers to evaluate both first and second derivatives at every iteration. Potentially the computation of these derivatives could require an overwhelming amount of work since for the Hessian $N^2$ cofactors are required. We show how the doubly stochastic structure and the properties of the objective may be exploited to calculate all cofactors from a single LU decomposition.

## 1 Introduction

The *Hamiltonian cycle problem* (HCP) is an important graph theory problem that features prominently in complexity theory because it is *NP-complete* [3]. HCP has also gained recognition because two special cases: the *Knight's tour* and the *Icosian game*, were solved by Euler and Hamilton, respectively. Finally, HCP is closely related to the well known Traveling Salesman problem.

The definition of HCP is the following: given a graph $\Gamma$ containing $N$ nodes, determine whether any simple cycles of length $N$ exist in the graph. These simple cycles of length $N$ are known as *Hamiltonian cycles*. If $\Gamma$ contains at least one Hamiltonian cycle (HC), we say that $\Gamma$ is a *Hamiltonian graph*. Otherwise, we say that $\Gamma$ is a *non-Hamiltonian graph*.

While there are many graph theory techniques that have been designed to solve HCP, another common approach is to associate a variable $x_{ij}$ with each arc $(i,j) \in \Gamma$, and solve an associated optimisation problem. A convenient method of representing these constraints is to use a matrix $P(x)$, where

$$p_{ij}(x) = \begin{cases} x_{ij}, & (i,j) \in \Gamma, \\ 0, & \text{otherwise.} \end{cases}$$

J.A. Filar
University of South Australia
E-mail: jerzy.filar@unisa.edu.au

M. Haythorpe
University of South Australia
E-mail: michael.haythorpe@unisa.edu.au

W. Murray
Stanford University
E-mail: walter@stanford.edu

The discrete nature of HCP naturally lends itself to integer programming optimisation problems. However, arising from an embedding of HCP in a Markov decision process (a practice initiated by Filar et al [2]), continuous optimisation problems that are equivalent to HCP have been discovered in recent times. In particular, it was demonstrated in [1] that if we define

$$A(P(x)) = I - P(x) + \frac{1}{N}ee^T, \tag{1}$$

where $e$ is a column vector with unit entries, then HCP is equivalent to solving the following optimisation problem:

$$\min \quad -\det(A(P(x)))$$

subject to

$$\sum_{j \in \mathcal{A}(i)} x_{ij} = 1, \quad i = 1, \ldots, N, \tag{2}$$

$$\sum_{i \in \mathcal{A}(j)} x_{ij} = 1, \quad j = 1, \ldots, N, \tag{3}$$

$$x_{ij} \geq 0, \quad \forall (i,j) \in \Gamma, \tag{4}$$

where $\mathcal{A}(i)$ is the set of nodes reachable in one step from node $i$. Constraints (2)–(4) are called the *doubly-stochastic* constraints. For neatness, we refer to constraints (2)–(4) as the set $\mathcal{DS}$, and we call the objective function $f(P(x))$. Then, the above problem can be represented as follows:

$$\min\{f(P(x)) \mid x \in \mathcal{DS}\}. \tag{5}$$

Note that we need to find a global minimiser and, typically, there are many of them. However, the number of global minimisers is typically extremely small compared to the number of local minimisers. One consequence of multiple global minimisers is that there is a similarly large number of stationary points. To distinguish such stationary points from minimisers, algorithms (see for example [4]) require the use of second derivatives. It is not hard to appreciate that evaluating such derivatives will be expensive even for moderately-sized problems unless some special structure is identified. We exploit the structure of the Hessian and the fact the points at which it is evaluated are in $\mathcal{DS}$. We assume an algorithm to solve this problem starts at a feasible point and all iterates remain feasible. For problems with nonlinear objectives this is almost always the best approach. Finding a point in $\mathcal{DS}$ is simple and does not involve evaluating the objective or its derivatives.

## 2 Preliminary results

Our work is based on some key properties of the LU factorisation of the matrix $I - P$. We adopt the following common notation. As noted we denote a vector of unit elements by $e$. In a given proof $e$ will still be used even when different occurrences may be of different dimension since its dimension can be inferred. The vector $e_j$ has all zero elements except the $j$th, which is unity. We will often denote the $(i,j)$th elements of a matrix, say $W$, by $w_{ij}$.

**Definition 1** A matrix $G$ is said to have property $S$ if

$$g_{ii} \geq 0, \quad \forall i = 1, \ldots, N,$$
$$g_{ij} \leq 0, \quad \forall i \neq j,$$
$$Ge = G^T e = 0.$$

**Theorem 1** *If $G$ has property $S$ the LU decomposition of $G$ exists.*

*Proof* Following common convention the diagonal elements of $L$ are 1. There is no loss of generality if we assume that $g_{11} \neq 0$. After one step of Gaussian elimination (GE) we obtain

$$G = \begin{pmatrix} 1 & 0^T \\ \ell & I \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & \hat{G} \end{pmatrix} \begin{pmatrix} u_{11} & u^T \\ 0 & I \end{pmatrix}.$$

Note that, expanding this matrix multiplication gives

$$G = \begin{pmatrix} u_{11} & u^T \\ u_{11}\ell & \hat{G} + \ell u^T \end{pmatrix},$$

and therefore $u^T = \begin{bmatrix} g_{12} & g_{13} & \cdots & g_{1N} \end{bmatrix}$ and $\ell = \dfrac{1}{u_{11}} \begin{bmatrix} g_{21} & g_{31} & \cdots & g_{N1} \end{bmatrix}^T$. Clearly, $u_{11} = g_{11} \geq 0$, and therefore $u^T \leq 0$ and $\ell \leq 0$. We have $Ge = 0$, which implies

$$\begin{pmatrix} 1 & 0^T \\ 0 & \hat{G} \end{pmatrix} \begin{pmatrix} u_{11} & u^T \\ 0 & I \end{pmatrix} e = 0.$$

It follows that $\hat{G}e = \hat{G}^T e = 0$. By definition we have

$$\hat{g}_{ij} = g_{i+1,j+1} - \ell_i u_j, \quad \forall i \neq j.$$

Since $l_i \leq 0$ and $u_j \leq 0$ it follows that $\hat{g}_{ij} \leq 0$, $\forall i \neq j$. From this result and $\hat{G}e = 0$ it follows that $\hat{g}_{ii} \geq 0$ and that $\hat{G}$ has property $S$. We can now proceed with the next step of GE. Note that if $\hat{g}_{11} = 0$ we must have the first row and column of $\hat{G}$ be zero, which implies the corresponding row of $U$ is zero and the off-diagonal elements of $L$ are zero. Consequently, we can proceed until the leading diagonal element of $\hat{G}$ is nonzero.

**Lemma 1** *Regardless of the rank of $G$ we have $u_{NN} = 0$.*

*Proof* After $N - 1$ steps of GE $\hat{G}$ is a $1 \times 1$ matrix. The only $1 \times 1$ matrix that satisfies property $S$ is 0.

**Corollary 1** *If $G$ satisfies property $S$ and $G = LU$, then $Ue = 0$.*

*Proof* The result follows immediately from the nonsingularity of $L$ and the property that $Ge = 0$.

**Lemma 2** *If $G$ satisfies property $S$ and $G = LU$, the off-diagonal elements of $L$ and $U$ are non-positive. Moreover, $-1 \leq l_{ij} \leq 0$ for $i \neq j$.*

*Proof* The first part follows immediately from the fact that $u^T \leq 0$ and $l \leq 0$ in each iteration of GE. Note that when performing the GE the pivot used is always the largest in magnitude. This is a consequence of $\hat{G}^T e = 0$ and the off-diagonal elements in the column of $\hat{G}$ being negative. Since the elements being eliminated are not bigger in magnitude than the pivot it follows that the magnitude of elements of $L$ are not bigger than 1.

The above result has important consequences for the condition number of $L$.

**Lemma 3** *If $G$ satisfies property $S$, is rank $N - 1$ and $G = LU$, then $L^T e = e_N$.*

*Proof* Since $U$ has rank $N - 1$ and $u_{NN} = 0$, the basis for the nullspace of $U^T$ is $\kappa e_N$, where $\kappa \neq 0$. From property $S$,

$$G^T e = U^T (L^T e) = 0$$

and hence we obtain $L^T e = \kappa e_N$, for some $\kappa$. Since $L$ has unit diagonal entries, we have $\kappa = 1$, which yields the result.

**Lemma 4** *When $G$ satisfies property $S$ and is rank $N - 1$ then $\bar{G} \equiv G + e e_N^T$ is nonsingular.*

*Proof* We have
$$\bar{G} = LU + e e_N^T = L(U + v e_N^T) = L\bar{U},$$

where $Lv = e$. Note that $\bar{U} = U + v e_N^T$ is upper triangular. Moreover, the $(N, N)$th element of $\bar{U}$ is $v_N$. Since
$$e^T L v = e_N^T v = v_N = e^T e = N,$$

it implies that $\bar{U}$, and hence $\bar{G}$, is nonsingular.

Clearly $G = I - P$ has property $S$. It should be noted that the existence of an LU decomposition for $G$ has been demonstrated previously for *irreducible* $P$ in Heyman [5]. In the context of this paper, we note that when $P(x)$ corresponds to the strict interior of $\mathcal{DS}$, irreducibility applies.

### 3 Computing $\det(A(P))$ using an LU decomposition

When appropriate we suppress the (fixed) argument $x$ in $P(x)$, $A(P(x))$ and $f(P(x))$, and write simply $P$, $A(P)$ (or just $A$) and $f(P)$, respectively. The most efficient way to compute the determinant of a matrix is to compute its LU decomposition. Normally we obtain $LU = \Pi A$, where $\Pi$ is a permutation matrix. However, we shall show that det $(A)$ may be computed from either the determinant of the leading principal minor of $I - P$ or of $\det(I - P + e e_N^T)$. Consequently, no permutation matrix is needed. This is of importance since unlike $A$ we expect $I - P$ to be sparse, which is the case of interest for finding HC. Factorising sparse matrices is much faster than factorising dense matrices. The difference is even greater when it is unnecessary to do numerical pivoting. Knowing apriori where the fill-in is in the factors will enable more efficient data management and avoid indirect addressing.

We assume an algorithm to solve (5) starts at a feasible point and all iterates remain feasible. For problems with nonlinear objectives this is almost always the best approach. Finding a point in $\mathcal{DS}$ is simple and does not involve evaluating the objective or its derivatives. It is known that the global minimiser of (5) is $-N$. Indeed, from Theorem 4.1, Proposition 4.4 and Proposition 4.6 in Ejov et al [1], we know that $-\det(A(P(x))) \in [-N, 0]$, where the upper bound is obtained when $A(P(x))$ is singular. However, we have no interest in finding the determinant of $A$ when it is singular. For the rest of this paper we assume $x \in \mathcal{DS}$ and that $A(P(x))$ is nonsingular. Since $A(P(x))$ is merely a rank-one correction of $I - P$, and the latter is always singular for $x \in \mathcal{DS}$, the nonsingularity of $A(P(x))$ implies that $I - P$ has rank $N - 1$.

3.1 Product forms of $A(P)$ and $\det(A(P))$

To calculate the objective function $f(P)$, its gradient and Hessian, we begin by performing an LU decomposition to obtain

$$LU = G = I - P. \tag{6}$$

While there may be a need to reorder the matrix to obtain a sparse factorisation. Since the pattern of nonzero elements in $G$ is symmetric then symmetric pivoting may be used, which is equivalent to renumbering the nodes of the graph. It follows we may assume the effort to compute the LU factors is $O(N^2)$. Note this reordering is done once since it depends only on the location of nonzeros and not their values. Whatever ordering is chosen the LU factors of $G$ reordered exist without the need for pivoting based on numerical considerations.

The outline of the derivation of $f(P)$ in terms of $U$ is as follows.

(1) We express A(P) as the product of three nonsingular factors.
(2) We show that two of these factors have a determinant of 1.
(3) We show that the third factor shares all but one eigenvalue with $U$, with the single different eigenvalue being $N$ (rather than 0).
(4) We express the determinant as a product of the first $N - 1$ diagonal elements of $U$, and $N$.

First, we express $A$ as a product of $L$ and another factor. Let $v$ be an $N \times 1$ vector, and $\bar{U}$ be an $N \times N$ matrix, such that

$$Lv = e, \quad \text{and} \quad \bar{U} := U + ve_N^T, \tag{7}$$

where $e_N^T = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$. Since $L$ is nonsingular, $v \neq 0$ exists and is unique, and therefore $\bar{U}$ is well-defined. The first $N-1$ columns of $\bar{U}$ are identical to those of $U$. Consequently $\bar{U}$ is also upper triangular. Since $LU = G$ satisfies property $S$, it follows from Lemma 1 that $Ue = 0$, and therefore $\bar{u}_{NN} = v_N = N$. Therefore,

$$\det \bar{U} = \prod_{i=1}^{N} \bar{u}_{ii} = N \prod_{i=1}^{N-1} u_{ii}, \tag{8}$$

where $u_{ii}$ is the $i$-th diagonal element of $U$. Exploiting (7) we may write

$$A = (I - P) + \frac{1}{N}ee^T \quad = \quad LU + \frac{1}{N}Lve^T \quad = \quad L(U + \frac{1}{N}ve^T)$$
$$= L(\bar{U} + v[\frac{1}{N}e^T - e_N^T]). \tag{9}$$

Since $\bar{U}$ is nonsingular we may define $w$ to be the unique solution to the system

$$\bar{U}^T w = \frac{1}{N}e - e_N. \tag{10}$$

Then, from (9)

$$A = L\left(\bar{U} + v\left[\frac{1}{N}e^T - e_N^T\right]\right) = L\left(\bar{U} + vw^T\bar{U}\right) = L\left(I + vw^T\right)\bar{U}. \tag{11}$$

We take the determinant of (11) to obtain

$$\det(A) = \det(L)\det(I + vw^T)\det(\bar{U}).\tag{12}$$

Note that, for any vectors $c$ and $d$,

$$\det(I + cd^T) = 1 + d^T c.\tag{13}$$

This is because $cd^T$ has one eigenvalue $d^T c$ of multiplicity 1 and an eigenvalue 0 of multiplicity $N - 1$. Consequently,

$$\det(I + vw^T) = 1 + w^T v,\tag{14}$$

which we substitute into expression (12) above.

**Lemma 5** *The inner-product $w^T v$ in (14) satisfies*

$$w^T v = 0.$$

*Proof* From their respective definitions (10) and (7),

$$w^T = \left(\frac{1}{N}e^T - e_N^T\right)(\bar{U})^{-1},\tag{15}$$

$$v = L^{-1}e.\tag{16}$$

Then, from (15)–(16) we obtain

$$w^T v = \left(\frac{1}{N}e^T - e_N^T\right)(\bar{U})^{-1}L^{-1}e \quad = \quad \left(\frac{1}{N}e^T - e_N^T\right)(L\bar{U})^{-1}e$$

$$= \left(\frac{1}{N}e^T - e_N^T\right)(I - P + ee_N^T)^{-1}e.\tag{17}$$

Since $I$, $P$ and $ee_N{}^T$ are all stochastic matrices, we know that $I - P + ee_N^T$ has row sums of 1 as well. Hence, its inverse also has row sums equal to 1, that is,

$$(I - P + ee_N^T)^{-1}e = e.\tag{18}$$

Substituting (18) into (17), we obtain

$$w^T v = \left(\frac{1}{N}e^T - e_N^T\right)e = 0,$$

which concludes the proof.

We now derive the main theorem of this subsection.

**Theorem 2** *Let $LU$ denote the LU decomposition of $I - P$ and $u_{ii}$ be the diagonal elements of $U$ then*

$$\det(A(P)) = N \prod_{i=1}^{N-1} u_{ii}.$$

*Proof* From (12), (14) and Lemma 5 we know that

$$\det(A(P)) = \det(L)(1+0)\det(\bar{U}).$$

From the construction of the LU decomposition we know that $\det(L) = 1$ and using (8) gives

$$\det(A(P)) = \det(\bar{U}) = N \prod_{i=1}^{N-1} u_{ii}.$$

This concludes the proof.

*Remark 1* Note that finding $v$ and $w$ is a simple process because $L$ and $\bar{U}^T$ are lower-triangular matrices, so we can solve the systems of linear equations in (10) and (7) directly.

3.2 Finding the gradient $g(P)$

Next we use the LU decomposition found in Subsection 3.1 to find the gradient of $f(P) = -\det(A(P))$. Note that since variables of $f(P)$ are entries $x_{ij}$ of the probability transition matrix $P(x)$, we derive an expression for $g_{ij}(P) := \frac{\partial f(P)}{\partial x_{ij}}$ for each $x_{ij}$ such that $(i,j) \in \Gamma$.

Consider vectors $a_j$ and $b_i$ satisfying the equations $\bar{U}^T a_j = e_j$ and $L b_i = e_i$, where $e_j$ is a zero vector except for a unit entry in the $j$-th column. Then, we define $Q := I - vw^T$, where $v$ and $w^T$ are as in (15)–(16). We prove the following result in this subsection:

$$g_{ij}(P) = \det(A(P))(a_j^T Q b_i), \tag{19}$$

where $g_{ij}(P)$ is the gradient vector element corresponding to the arc $(i,j) \in \Gamma$.

The outline of the derivation of (19) is as follows.

(1) We represent each element $g_{ij}(P)$ of the gradient vector as a cofactor of $A(P)$.
(2) We construct an elementary matrix that transforms matrix $A(P)$ into a matrix with determinant equal to the above cofactor of $A(P)$.
(3) We then express the element $g_{ij}(P)$ of the gradient vector as the product of $\det(A(P))$ and the determinant of the elementary matrix, the latter of which is shown to be equal to $a_j^T Q b_i$.

For any matrix $V = (v_{ij})_{i,j=1}^{N,N}$ it is well-known (e.g., see May [6]) that $\frac{\partial \det(V)}{\partial v_{ij}} = (-1)^{i+j} \det(V^{ij})$, where $V^{ij}$ is the $(i,j)$-th minor of $V$. That is, $\frac{\partial \det(V)}{\partial v_{ij}}$ is the $(i,j)$-th cofactor of $V$. Since the $(i,j)$-th entry of $A(P)$ is simply $a_{ij} = \delta_{ij} - x_{ij} + \frac{1}{N}$ (where $\delta_{ij}$ is the Kronecker delta that is 1 if $i = j$, and 0 otherwise), it now follows that

$$g_{ij}(P) = \frac{\partial f(P)}{\partial x_{ij}} = \frac{\partial \left[ -\det A(P) \right]}{\partial a_{ij}} \frac{d a_{ij}}{d x_{ij}} = (-1)^{i+j} \det\left( A^{ij}(P) \right). \tag{20}$$

However, rather than finding the cofactor we calculate (20) by finding the determinant of a modification of $A$ in which the $i$th row has been replaced with $e_j^T$. Since $A$ is a full-rank matrix, it is possible to perform

row operations to achieve this. Suppose $A$ is composed of rows $r_1^T$, $r_2^T$, ..., $r_N^T$. Then, we perform the following row operation:

$$r_i^T \rightarrow \alpha_j(1)r_1^T + \alpha_j(2)r_2^T + \ldots + \alpha_j(N)r_N^T, \tag{21}$$

where $\alpha_j(i)$ is the $i$-th element of vector $\alpha_j$ and $A^T\alpha_j = e_j$. This row operation replaces the $i$-th row of $A$ with $\alpha_j^T A = e_j^T$, as desired.

In this case, from (11), $A^T = \bar{U}^T(I + wv^T)L^T$. Since $A$ is nonsingular $\alpha_j$ can be found directly:

$$\begin{aligned}
\alpha_j = \left(A^T\right)^{-1} e_j &= \left[\bar{U}^T(I + wv^T)L^T\right]^{-1} e_j \\
&= (L^T)^{-1}(I + wv^T)^{-1}(\bar{U}^T)^{-1}e_j.
\end{aligned} \tag{22}$$

**Lemma 6**

$$\left(I + wv^T\right)^{-1} = I - wv^T.$$

*Proof* Consider

$$\begin{aligned}
(I + wv^T)(I - wv^T) &= I - wv^T + wv^T - wv^T wv^T \\
&= I - wv^T wv^T \\
&= I, \text{ because } v^T w = w^T v = 0, \qquad \text{from Lemma 5.}
\end{aligned}$$

Therefore, $\left(I + wv^T\right)^{-1} = \left(I - wv^T\right).$

Taking the above result and substituting into (22), we obtain

$$\alpha_j = (L^T)^{-1}(I - wv^T)(\bar{U}^T)^{-1}e_j. \tag{23}$$

Next, we define an elementary matrix $E_{ij}$ by

$$E_{ij} := I - e_i e_i^T + e_i \alpha_j^T, \tag{24}$$

and note that it performs the desired row operation (21) on A because

$$E_{ij}A = A - e_i r_i^T + e_i e_j^T,$$

in effect replacing the $i$-th row of $A$ with $e_j^T$. Therefore,

$$g_{ij}(P) = (-1)^{i+j} \det\left(A^{ij}\right) = \det(E_{ij}A) = \det(E_{ij})\det(A). \tag{25}$$

From (24), we rewrite $E_{ij} = I - e_i(e_i - \alpha_j)^T$. Then, from (13) we obtain

$$\det(E_{ij}) = 1 - (e_i - \alpha_j)^T e_i = 1 - e_i^T e_i + \alpha_j^T e_i = \alpha_j^T e_i. \tag{26}$$

Substituting (23) into (26) we obtain

$$\det(E_{ij}) = e_j^T(\bar{U})^{-1}(I - vw^T)(L)^{-1}e_i. \tag{27}$$

For convenience we define $Q := I - vw^T$. Then

$$\det(E_{ij}) = a_j^T Q b_i, \quad \text{where } \bar{U}^T a_j = e_j \text{ and } L b_i = e_i. \tag{28}$$

We now derive the main result of this subsection.

**Proposition 1** *The general gradient element of $f(P)$ is given by*

$$g_{ij}(P) = \frac{\partial f(P)}{\partial x_{ij}} = \det(A(P))(a_j^T Q b_i). \tag{29}$$

*Proof* Substituting (28) into (25) immediately yields the result.

*Remark 2* Note that we can calculate all $a_j$ and $b_i$ in advance, by solving the systems of linear equations in (28), again in reduced row echelon form. Then, for the sake of efficiency we first calculate

$$\hat{q}_j^T := a_j^T Q, \quad j = 1, \dots, N, \tag{30}$$

and then calculate

$$\hat{q}_{ij} := \hat{q}_j^T b_i, \quad i = 1, \dots, N, \quad j = 1, \dots, N. \tag{31}$$

This allows us to rewrite the formula for $g_{ij}(P)$ as

$$g_{ij}(P) = -f(P)\hat{q}_{ij}.$$

3.3 Finding the Hessian matrix $H(P)$

Here, we show that the LU decomposition found in Subsection 3.1 can also be used to calculate the Hessian of $f(P)$ efficiently. Consider $g_{ij}$ and $\hat{q}_{ij}$ as defined in (29) and (31) respectively. We prove the following result in this subsection:

$$H_{[ij],[k\ell]}(P) := \frac{\partial^2 f(P)}{\partial x_{ij} \partial x_{k\ell}} = g_{kj}\hat{q}_{i\ell} - g_{ij}\hat{q}_{k\ell},$$

where $H_{[ij],[k\ell]}$ is the general element of the Hessian matrix corresponding to arcs $(i,j)$ and $(k,\ell) \in \Gamma$.

The outline of the derivation is as follows.

(1) We represent each element $H_{[ij],[k\ell]}(P)$ of the Hessian matrix as a cofactor of a minor of $A(P)$.
(2) We construct a second elementary matrix that in conjunction with $E_{ij}$ (see (24)) transforms matrix $A(P)$ into one with a determinant equivalent to the $(k,\ell)$-th cofactor of $A^{ij}(P)$.
(3) We then show that the general element of the Hessian matrix is the product of $\det(A(P))$ and the determinants of the two elementary matrices.
(4) Using results obtained from finding $g(P)$ in Subsection 3.2, we obtain these values immediately.

We define $A^{[ij],[k\ell]}$ to be the matrix $A$ with rows $i,k$ and columns $j,\ell$ removed. An argument similar to that for $g_{ij}(P)$ in the previous subsection can be made that finding

$$H_{[ij],[k\ell]}(P) = \frac{\partial^2 f(P)}{\partial x_{ij} \partial x_{k\ell}} = (-1)^{(i+j+k+\ell+1)} \det(A^{[ij],[k\ell]}), \ i \neq k, \ j \neq \ell, \tag{32}$$

is equivalent to finding the negative determinant of $A$ with the $i$th and $k$th rows changed to $e_j^T$ and $e_\ell^T$ respectively. That is,

$$\frac{\partial^2 f(P)}{\partial x_{ij} \partial x_{k\ell}} = -\det(\hat{E}_{k\ell} E_{ij} A(P))$$

$$= -\det(\hat{E}_{k\ell}) \det(E_{ij}) \det(A(P)), \tag{33}$$

where $\hat{E}_{k\ell}$ is an additional row operation constructed to change row $k$ of $E_{ij}A$ into $e_\ell^T$. Note that if $i = k$ or $j = \ell$, the matrix $A^{[ij],[k\ell]}$ is no longer square and the determinant no longer exists. If this occurs, we define $H_{[ij],[k\ell]} := 0$. If both $i = k$ and $j = \ell$, we also define $H_{[ij],[k\ell]} := 0$, as the determinant is linear in each element of $A(P)$.

Consider $E_{ij}A$ composed of rows $\hat{r}_1^T$, $\hat{r}_2^T$, ..., $\hat{r}_N^T$. Then, we perform the following row operation:

$$\hat{r}_k \rightarrow \gamma_\ell(1)\hat{r}_1 + \gamma_\ell(2)\hat{r}_2 + \ldots + \gamma_\ell(N)\hat{r}_N, \tag{34}$$

where $(E_{ij}A)^T \gamma_\ell = e_\ell$. Then, similarly to (22), we directly find $\gamma_\ell$:

$$\gamma_\ell = (E_{ij}^T)^{-1}(L^T)^{-1}(I - wv^T)(\overline{U}^T)^{-1}e_\ell. \tag{35}$$

Next, in a similar fashion to (24), we construct an elementary matrix $\hat{E}_{k\ell}$

$$\begin{aligned}
\hat{E}_{k\ell} &= I - e_k e_k^T + e_k \gamma_\ell^T \\
&= I - e_k(e_k^T - \gamma_\ell^T).
\end{aligned} \tag{36}$$

Then, we evaluate $\det(\hat{E}_{k\ell})$:

$$\begin{aligned}
\det(\hat{E}_{k\ell}) &= 1 - (e_k^T - \gamma_\ell^T)e_k \\
&= 1 - 1 + \gamma_\ell^T e_k \\
&= e_\ell^T (\overline{U})^{-1} Q L^{-1} (E_{ij})^{-1} e_k.
\end{aligned} \tag{37}$$

Recall from (24) that $E_{ij} = I - e_i(e_i^T - \alpha_j^T)$. We have

$$(E_{ij})^{-1} = I + \frac{1}{\alpha_j^T e_i} e_i(e_i^T - \alpha_j^T). \tag{38}$$

Recall from (25) that $g_{ij} = \det(A)\det(E_{ij})$, and from (26) that $\alpha_j^T e_i = \det(E_{ij}) \neq 0$, and therefore (38) holds. Then,

$$\alpha_j^T e_i = \frac{g_{ij}}{\det(A)}. \tag{39}$$

Substituting (39) into (38) we obtain

$$(E_{ij})^{-1} = I + \frac{\det(A)}{g_{ij}} e_i(e_i^T - \alpha_j^T), \tag{40}$$

and further substituting (40) into (37), we obtain

$$\begin{aligned}
\det(\hat{E}_{k\ell}) &= e_\ell^T (\overline{U})^{-1} Q L^{-1} \left( I + \frac{\det(A)}{g_{ij}} e_i(e_i^T - \alpha_j^T) \right) e_k \\
&= e_\ell^T (\overline{U})^{-1} Q L^{-1} \left( e_k + \frac{\det(A)}{g_{ij}} e_i e_i^T e_k - \frac{\det(A)}{g_{ij}} e_i \alpha_j^T e_k \right).
\end{aligned} \tag{41}$$

Note that since $i \neq k$, $e_i e_i^T e_k = 0$, and from (39), $\alpha_j^T e_k = \frac{g_{kj}}{\det(A)}$. Hence, from (41) and (28) we obtain

$$\begin{aligned}
\det(\hat{E}_{k\ell}) &= e_\ell^T (\overline{U})^{-1} Q L^{-1} (e_k - \frac{g_{kj}}{g_{ij}} e_i) \\
&= a_\ell^T Q(b_k - \frac{g_{kj}}{g_{ij}} b_i).
\end{aligned} \tag{42}$$

We now derive the main result of this subsection.

**Proposition 2** *The general element of the Hessian of $f(P)$ is given by*

$$H_{[ij],[k\ell]} = g_{kj}\hat{q}_{i\ell} - g_{ij}\hat{q}_{k\ell},$$

*where $\hat{q}_{i\ell}$ and $\hat{q}_{k\ell}$ are defined in (31).*

*Proof* From (33) and (25), we can see that $H_{[ij],[k\ell]} = -\det(\hat{E}_{k\ell})g_{ij}$. Then, from (42), $\det(\hat{E}_{k\ell}) = a_\ell^T Q(b_k - \frac{g_{kj}}{g_{ij}}b_i)$ and so $H_{[ij],[k\ell]} = -a_\ell^T Q(b_k g_{ij} - b_i g_{kj})$.

In order to improve computation time, we take advantage of the fact that we evaluate every $\hat{q}_{ij}$ while calculating the gradient to rewrite the second order partial derivatives of $f(P)$ as

$$\begin{aligned} H_{[ij],[k\ell]} &= g_{kj}a_\ell^T Qb_i - g_{ij}a_\ell^T Qb_k \\ &= g_{kj}\hat{q}_{i\ell} - g_{ij}\hat{q}_{k\ell}. \end{aligned} \tag{43}$$

This concludes the proof.

*Remark 3* Note that in practice, we do not calculate some $g_{kj}$'s when calculating $g(P)$ as an arc $(k,j)$ need not exist in the graph. In these cases we find $g_{jk}$ using the gradient formula, $g_{jk} = -f(P)(\hat{q}_{jk})$, which remains valid despite arc $(k,j)$ not appearing in the graph.

3.4 Leading principal minor

It is, perhaps, interesting that instead of using the objective function $f(P) = -\det\left(I - P + \frac{1}{N}ee^T\right)$, it is also possible to use $f^1(P) := -\det(G^{NN}(P))$, the negative determinant of the leading principal minor of $I - P$. The following, somewhat surprising, result justifies this claim.

**Theorem 3** *(1) $f^1(P) = \frac{1}{N}f(P) = -\frac{1}{N}\det\left(I - P + \frac{1}{N}ee^T\right)$.*
*(2) If the graph is Hamiltonian, then*

$$\min_{P \in \mathcal{DS}} f^1(P) = -1. \tag{44}$$

*Proof* First, we show part (1), that is, $f^1(P) = \frac{1}{N}f(P)$. To find $f^1(P)$, we construct $LU = I - P$ as before, and define $\hat{L}$, $\hat{U}$ as:

$$\hat{L} = \begin{bmatrix} e_1^T L \\ \vdots \\ e_{N-1}^T L \\ e_N^T \end{bmatrix}, \quad \hat{U} = \begin{bmatrix} Ue_1 & \cdots & Ue_{N-1} & e_N \end{bmatrix}. \tag{45}$$

That is, $\hat{L}$ is the same as $L$ with the last row replaced by $e_N^T$, and $\hat{U}$ is the same as $U$ with the last column replaced with $e_N$. Then consider

$$\hat{L}\hat{U} = \begin{bmatrix} e_1^T L \\ \vdots \\ e_{N-1}^T L \\ e_N^T \end{bmatrix} \begin{bmatrix} Ue_1 & \cdots & Ue_{N-1} & e_N \end{bmatrix}$$

$$= \begin{bmatrix} e_1^T L U e_1 & \ddots & e_1^T L U e_{N-1} & e_1^T L e_N \\ \vdots & \ddots & \vdots & \vdots \\ e_{N-1}^T L U e_1 & \cdots & e_{N-1}^T L U e_{N-1} & e_{N-1}^T L e_N \\ e_N^T U e_1 & \cdots & e_N^T U e_{N-1} & e_N^T e_N \end{bmatrix}.$$

Since $L$ is lower-triangular, $e_i^T L e_N = 0$ for all $i \neq N$. Likewise, since $U$ is upper-triangular, $e_N^T U e_j = 0$ for all $j \neq N$. Therefore the above matrix simplifies to

$$\hat{L}\hat{U} = \begin{bmatrix} e_1^T L U e_1 & \cdots & e_1^T L U e_{N-1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ e_{N-1}^T L U e_1 & \cdots & e_{N-1}^T L U e_{N-1} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix},$$

which is the same as $LU$ with the bottom row and rightmost column removed, and a 1 placed in the bottom-right element. Therefore, $\det(\hat{L}\hat{U}) = \det(G^{NN}(P))$, and consequently

$$f^1(P) = -\det(\hat{L})\det(\hat{U}). \tag{46}$$

Note that $\hat{L}$ and $\hat{U}$ are triangular matrices, so

$$\det\left(\hat{L}\right) = \prod_{i=1}^{N} \hat{l}_{ii}, \qquad \text{and} \qquad \det\left(\hat{U}\right) = \prod_{i=1}^{N} \hat{u}_{ii}.$$

However, only the last diagonal elements of $\hat{L}$ and $\hat{U}$ are different from $L$ and $\bar{U}$ (see (7)) respectively, so

$$\det\left(\hat{L}\right) = \hat{l}_{NN} \prod_{i=1}^{N-1} l_{ii}, \qquad \text{and} \qquad \det\left(\hat{U}\right) = \hat{u}_{NN} \prod_{i=1}^{N-1} \bar{u}_{ii}. \tag{47}$$

Now, since $\hat{l}_{NN} = l_{NN} = 1$, we have

$$\det\left(\hat{L}\right) = \det(L) = 1. \tag{48}$$

We also have $\hat{u}_{NN} = 1$, but by Lemma 4, $\bar{u}_{NN} = N$ and hence

$$\det\left(\hat{U}\right) = \frac{1}{N} \det\left(\bar{U}\right). \tag{49}$$

Therefore, substituting (48) and (49) into (46) we obtain

$$f^1(P) = -\det\left(\hat{L}\right)\det\left(\hat{U}\right)$$

$$= -\frac{1}{N}\det\left(\bar{U}\right)$$

$$= -\frac{1}{N}\det\left(I - P + \frac{1}{N}ee^T\right) = \frac{1}{N}f(P).$$

Therefore, part (1) is proved.

The proof of part (2) of Theorem 3 follows directly from the fact that $\min f(P) = -N$ (proved in [1]), and part (1).

*Remark 4* Using the leading principal minor has the advantage that the rank-one modification $\frac{1}{N}ee^T$ is not required, which makes calculating the gradient and the Hessian even simpler than described in Subsection 3.2 and Subsection 3.3 respectively. The derivation of the gradient and Hessian formulae for the negative determinant of the leading principal minor follows the same process as that for the determinant function, except that the matrix $Q = I - vw^T$ is not required.

The formulae for $f^1(P)$, $g^1(P)$ and $H^1(P)$ then reduce to

$$f^1 = -\prod_{i=1}^{N-1} u_{ii}, \tag{50}$$

$$g_{ij}^1 = -f^1(P)(a_j^1)^T b_i^1, \tag{51}$$

$$H_{[i,j],[k,\ell]}^1 = g_{kj}^1 (a_\ell^1)^T b_i^1 - g_{ij}^1 (a_\ell^1)^T b_k^1, \tag{52}$$

where

$$\hat{L} b_i^1 = e_i, \tag{53}$$

$$\hat{U}^T a_j^1 = e_j. \tag{54}$$

*Remark 5* In practice, the determinant of the leading principal minor is used rather than that of the whole matrix. It is simpler, more efficient and the optimal value is independent of the graph. It eliminates the need to scale any parameters by the size of the graph. When $f^1(P)$ is used in lieu of $f(P)$ the corresponding gradient vector and Hessian matrix are denoted by $g^1(P)$ and $H^1(P)$, respectively.

## 4 LU decomposition-based evaluation algorithm

The algorithm for computing $f^1(P)$, $g^1(P)$, $H^1(P)$ is given here, along with the complexity of each step of the algorithm. Let $k$ denote the average degree of the graph, that is, there are $kN$ edges.

| | Complexity |
|---|---|
| **Input**: $P$<br>**Output**: $f^1(P), g^1(P), H^1(P)$ | |
| **begin** | |
| 1) Perform LU decomposition to find $LU = I - P$. | $O(kN^2)$ |
| 2) Calculate $\hat{L}$ and $\hat{U}$, using (45). | $O(N)$ |
| 3) Calculate each $(a_j^1)^T$ and $b_i^1$, using (53) and (54). | $O(N^3)$ |
| 4) Calculate each $(a_j^1)^T b_i^1$. | $O(N^3)$ |
| 5) Calculate $f^1(P) = -\prod_{i=1}^{N-1} u_{ii}$. | $O(N)$ |
| 6) Calculate each $g_{ij}^1(P) = -f^1(P)(a_j^1)^T b_i^1$. | $O(kN)$ |
| 7) Calculate each $H_{[ij],[k\ell]}^1(P) = \begin{cases} g_{kj}^1(a_\ell^1)^T b_i^1 - g_{ij}^1(a_\ell^1)^T b_k^1, & i \neq k \text{ and } j \neq \ell \text{ and } i,j,k,l \neq N \\ 0, & \text{otherwise.} \end{cases}$ | $O(k^2 N^2)$ |
| **end** | |

Function evaluations algorithm

If the graph is sparse, the complexity of the above algorithm is $O(N^3)$. However, for sufficiently dense graphs (that is, $k > \sqrt{N}$) the complexity of the above algorithm is $O(k^2 N^2)$. Note that each element of the Hessian is calculated in $O(1)$ time, because they simply involve scalar multiplication where all of the scalars have already been calculated in earlier steps, that is, the gradient terms in step 6, and each $\left(a_i^1\right)^T b_l^1$ in step 4.

These bounds are considerably better than the $O(k^3 N^4)$ bound that applies if we simply perform an LU decomposition for each element in the Hessian and gradient.

*Example 1* Consider the following six-node cubic graph $\Gamma_6$.
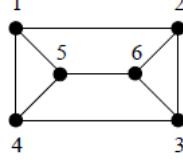


The adjacency matrix of $\Gamma_6$ is

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Consider a point $x$ such that,

$$P(x) = \begin{bmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{2}{3} & 0 & \frac{1}{6} & 0 & 0 & \frac{1}{6} \\ 0 & \frac{1}{6} & 0 & \frac{2}{3} & 0 & \frac{1}{6} \\ \frac{1}{6} & 0 & \frac{2}{3} & 0 & \frac{1}{6} & 0 \\ \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 & \frac{2}{3} \\ 0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{2}{3} & 0 \end{bmatrix}.$$

Performing the LU decomposition of $I - P$ using MATLAB's `lu` routine we obtain matrices $L$ and $U$ (given to four decimal places)

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.6667 & 1 & 0 & 0 & 0 & 0 \\ 0 & -0.3000 & 1 & 0 & 0 & 0 \\ -0.1667 & -0.2000 & -0.7368 & 1 & 0 & 0 \\ -0.1667 & -0.2000 & -0.0351 & -0.5556 & 1 & 0 \\ 0 & -0.3000 & -0.2281 & -0.4444 & -1.0000 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & -0.6667 & 0 & -0.1667 & -0.1667 & 0 \\ 0 & 0.5556 & -0.1667 & -0.1111 & -0.1111 & -0.1667 \\ 0 & 0 & 0.9500 & -0.7000 & -0.0333 & -0.2167 \\ 0 & 0 & 0 & 0.4342 & -0.2412 & -0.1930 \\ 0 & 0 & 0 & 0 & 0.8148 & -0.8148 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consequently, $\hat{L} = \begin{bmatrix} L^T e_1 & \cdots & L^T e_{N-1} & e_N \end{bmatrix}^T$ and $\hat{U} = \begin{bmatrix} U e_1 & \cdots & U e_{N-1} & e_N \end{bmatrix}$ are simply

$$\hat{L} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.6667 & 1 & 0 & 0 & 0 & 0 \\ 0 & -0.3000 & 1 & 0 & 0 & 0 \\ -0.1667 & -0.2000 & -0.7368 & 1 & 0 & 0 \\ -0.1667 & -0.2000 & -0.0351 & -0.5556 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \hat{U} = \begin{bmatrix} 1 & -0.6667 & 0 & -0.1667 & -0.1667 & 0 \\ 0 & 0.5556 & -0.1667 & -0.1111 & -0.1111 & 0 \\ 0 & 0 & 0.9500 & -0.7000 & -0.0333 & 0 \\ 0 & 0 & 0 & 0.4342 & -0.2412 & 0 \\ 0 & 0 & 0 & 0 & 0.8148 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

For all $i, j$, we calculate the $a_j^1$ and $b_i^1$ vectors using (53) and (54). Namely,

$$a_1^1 = \begin{bmatrix} 1 \\ 1.2 \\ 0.2105 \\ 1.0303 \\ 0.6818 \\ 0 \end{bmatrix}, a_2^1 = \begin{bmatrix} 0 \\ 1.8 \\ 0.3158 \\ 0.9697 \\ 0.5455 \\ 0 \end{bmatrix}, a_3^1 = \begin{bmatrix} 0 \\ 0 \\ 1.0526 \\ 1.6970 \\ 0.5455 \\ 0 \end{bmatrix}, a_4^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2.3030 \\ 0.6818 \\ 0 \end{bmatrix}, a_5^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1.2273 \\ 0 \end{bmatrix}, a_6^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

$$b_1^1 = \begin{bmatrix} 1 \\ 0.6667 \\ 0.2 \\ 0.4474 \\ 0.5556 \\ 0 \end{bmatrix}, b_2^1 = \begin{bmatrix} 0 \\ 1 \\ 0.3 \\ 0.4211 \\ 0.4444 \\ 0 \end{bmatrix}, b_3^1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0.7368 \\ 0.4444 \\ 0 \end{bmatrix}, b_4^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0.5556 \\ 0 \end{bmatrix}, b_5^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, b_6^1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

We can now represent each $(a_j^1)^T b_i^1$ as the $ij$-th element of the matrix

$$\left[(a_j^1)^T (b_i^1)\right]_{i,j=1}^{N,N} = \left[ \begin{array}{cccccc} 2.6818 & 2 & 1.2727 & 1.4091 & 0.6818 & 0 \\ 2 & 2.5455 & 1.2727 & 1.2727 & 0.5455 & 0 \\ 1.2727 & 1.2727 & 2.5455 & 2 & 0.5455 & 0 \\ 1.4091 & 1.2727 & 2 & 2.6818 & 0.6818 & 0 \\ 0.6818 & 0.5455 & 0.5455 & 0.6818 & 1.2273 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right].$$

Then, $f^1(P) = \prod_{i=1}^{N-1} \hat{u}_{ii} \approx -0.1867$. Note that we can directly verify the preceding by confirming that $\det(A(P)) \approx 1.1204 = -6(f^1(P))$.

The gradient vector is then found using (51). Note that we are only interested in the gradient elements for the eighteen arcs in the graph; this yields, to three decimal places:

$$g^1(P) \approx \left[ \begin{array}{cccccccccccccccccc} 0.374 & 0.263 & 0.127 & 0.374 & 0.238 & 0 & 0.238 & 0.374 & 0 & 0.263 & 0.374 & 0.127 & 0.127 & 0.127 & 0 & 0 & 0 & 0 \end{array} \right].$$

Finally, the Hessian is found using (52), given here to two decimal places:

$$H^1(P) \approx \left[ \begin{array}{cccccccccccccccccc} 0 & 0 & 0 & 0.53 & 0.13 & 0 & 0 & -0.41 & 0 & 0.11 & -0.44 & -0.09 & 0.02 & -0.11 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.11 & -0.03 & 0 & 0.41 & 0 & 0 & 0.97 & 0.11 & 0.16 & 0.16 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.02 & -0.03 & 0 & 0.04 & -0.11 & 0 & 0.16 & -0.09 & 0 & 0.53 & 0.24 & 0 & 0 & 0 & 0 \\ 0.53 & 0.11 & 0.02 & 0 & 0 & 0 & 0.13 & -0.44 & 0 & -0.41 & -0.11 & 0 & -0.09 & 0 & 0 & 0 & 0 & 0 \\ 0.13 & -0.03 & -0.03 & 0 & 0 & 0 & 0.91 & 0.13 & 0 & 0.41 & 0 & 0.04 & 0.04 & -0.03 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.41 & 0.04 & 0.13 & 0.91 & 0 & 0 & 0 & 0 & -0.03 & 0.13 & -0.03 & -0.03 & 0.04 & 0 & 0 & 0 & 0 \\ -0.41 & 0 & -0.11 & -0.44 & 0.13 & 0 & 0 & 0 & 0 & 0.11 & 0.53 & 0.02 & -0.09 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.11 & 0.97 & 0.16 & 0 & 0.41 & 0 & -0.03 & 0.11 & 0 & 0 & 0 & 0 & 0 & 0.16 & 0 & 0 & 0 & 0 \\ -0.44 & 0.11 & -0.09 & -0.41 & 0 & 0 & 0.13 & 0.53 & 0 & 0 & 0 & 0 & -0.11 & 0.02 & 0 & 0 & 0 & 0 \\ -0.09 & 0.16 & 0 & -0.11 & 0.04 & 0 & -0.03 & 0.02 & 0 & 0 & 0 & 0 & 0.24 & 0.53 & 0 & 0 & 0 & 0 \\ 0.02 & 0.16 & 0.53 & 0 & 0.04 & 0 & -0.03 & -0.09 & 0 & -0.11 & 0.24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.11 & 0 & 0.24 & -0.09 & -0.03 & 0 & 0.04 & 0 & 0 & 0.16 & 0.02 & 0.53 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right].$$

## References

1. V. Ejov, J. Filar, W. Murray, and G. T. Nguyen. Determinants and longest cycles of graphs. *SIAM Journal on Discrete Mathematics*, 22(3):1215–1225, 2009.
2. J. A. Filar and D. Krass. Hamiltonian cycles and Markov chains. *Mathematics of Operations Research*, 19:223–237, 1994.
3. M. R. Garey, D. S. Johnson, and R. E. Tarjan. The planar Hamiltonian circuit problem is NP-complete. *SIAM Journal on Computing*, 5(4):704–714, 1976.
4. M. Haythorpe. *Markov Chain Based Algorithms for the Hamiltonian Cycle Problem*. PhD thesis, University of South Australia, 2010. Available at: http://www.stanford.edu/group/SOL/dissertations/michael-haythorpe-thesis.pdf.
5. D. Heyman. A decomposition theorem for infinite stochastic matrices. *Journal of Applied Probability*, 32:893–901, 1995.
6. K. May. Derivatives of Determinants and Other Multilinear Functions. *Mathematics Magazine*, 38(5):307–308, 1965.