ELSEVIER Linear Algebra and its Applications 321 (2000) 95–112
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Eigenfunction properties and approximations of selected incidence matrices employed in spatial analyses

Daniel A. Griffith [1]

Department of Geography and Interdisciplinary Statistics Program, 144 Eggers Hall, Syracuse University, Syracuse, NY 13244-1090, USA

Received 7 November 1998; accepted 5 January 2000

Submitted by G.P.H. Styan

Abstract

Mathematical properties of eigenfunctions of selected incidence matrices appearing in spatial statistics formulae are summarized. Seven theorems are proposed and proved, and three conjectures are posited. Results summarized here allow the determinant of massively large $n \times n$ geographic weights matrices to be accurately approximated. In addition, the behavior of eigenfunctions for graphs affiliated with a linear configuration of connected nodes are better understood. © 2000 Elsevier Science Inc. All rights reserved.

Keywords: Eigenfunction; Incidence matrix; Spatial analysis; Square tessellation; Hexagonal tessellation; Stochastic matrix

1. Introduction

Eigenfunction research, whose results are summarized here was inspired by several common linear algebra problems associated with massively large georeferenced data sets, or data sets that are locationally tagged to the earth's surface [10,27]. Two such types of data sets to which this paper is relevant are those generated by satellite remote sensing, and those constructed under the auspices of the US Environmental

E-mail address: griffith@maxwell.syr.edu (D.A. Griffith).

[1] Part of this research was supported by the National Science Foundation, research grant #SES-91-22232.

PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 0 3 1 - 8

Protection Agency's Environmental Monitoring and Assessment Program (US EPA EMAP; [26]). The geographic arrangement of locational tagging in this first case is the regular square tessellation, and in this second case is the regular hexagonal tessellation [2, p. 7]. The duals of both of these geographic configurations can be analyzed as graphs, and depicted with an incidence or binary 0–1 matrix $\mathbf{C}$ whose c_{ij} element value is 1 if locations i and j are nearby, and 0 otherwise. Throughout this paper definition of a graph link is based upon the rook's definition of adjacency, drawing upon an analogy with chess moves. Accordingly, nearby is defined as whether or not a link exists between the nodes of a tessellation dual. This definition results in matrix $\mathbf{C}$ being symmetric, with all diagonal entries of 0. Historically matrix theory and linear algebra have been used extensively to analyze such adjacency matrices of graphs. Furthermore, Chung [8, p. 135] notes that "[a] crucial part of spectral graph theory concerns understanding the behavior of the eigenfunctions" of these matrices. Meanwhile, the most interesting linear algebra problem solution contributed to by results reported in this paper is the estimation of eigenfunctions, and hence also matrix determinants for large order sparse matrices. While numerical results for such problems are topical (e.g., [3,5]), a number of analytical results are reported in this paper.

1.1. Basic definitions and terminology

Spatial statistics is a very specialized subdiscipline, with its geographic applications involving terminology unfamiliar to many scholars. The interested reader may wish to consult Cressie [10], Griffith and Layne [17], and Anselin [1], among others, for different overview perspectives of this subdiscipline. The focus in this paper is on specific matrices commonly used in spatial statistics work. These include not only the previously defined $n \times n$ incidence matrix $\mathbf{C}$, but also $n \times n$ matrices $\mathbf{W}$, the popular row-standardized, stochastic version of matrix $\mathbf{C}$ (i.e., $w_{ij} = c_{ij}/\sum_{j=1}^{n} c_{ij}$), and

$$\left(\mathbf{I} - \frac{\mathbf{11}^{\mathrm{T}}}{n}\right)\mathbf{C}\left(\mathbf{I} - \frac{\mathbf{11}^{\mathrm{T}}}{n}\right), \tag{1}$$

with $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$ being the idempotent projection matrix commonly found in conventional statistics [21, p. 115], where $\mathbf{I}$ is the identity matrix, $\mathbf{1}$ is an $n \times 1$ vector of ones, and the superscript T denotes matrix transpose.

The importance of these three matrices in spatial statistics lies in their appearances in computational formulae. Their eigenfunctions also play important roles in spatial statistical theory.

A common optimization problem in spatial statistics, involving matrices $\mathbf{C}$ and $\mathbf{W}$ and their eigenfunctions, is the minimization of the following function, which actually is a log-likelihood function

$$\text{constant} - \frac{n}{2}\mathrm{LN}(\sigma^2) + \frac{1}{2}\mathrm{LN}[\det(\mathbf{V})] - (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{V}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})/(2\sigma^2),$$

where $\det(\cdot)$ denotes the operation of matrix determinant, LN denotes the natural logarithm, $\mathbf{Y}$ is an $n \times 1$ vector of data values, $\mathbf{X}$ is an $n \times p$ matrix of data for p different additional variables, $\mathbf{X}\boldsymbol{\beta}$ and σ, respectively, denote the standard regression mean and variance parameters of linear statistical analysis, and matrix $\mathbf{V}$ is a function of the connectivity matrix $\mathbf{C}$ or $\mathbf{W}$. The interested reader can find out more about this expression by consulting a multivariate textbook such as [21, pp. 143–150]. From calculus one finds that the term $\frac{1}{2}\mathrm{LN}[\det(\mathbf{V})]$ is a Jacobian term. Ord [28] shows that it can be written as a function of the eigenvalues of matrix $\mathbf{C}$ or $\mathbf{W}$, depending upon which matrix is used to define $\mathbf{V}$, and derives the n analytical eigenvalues, λ_{kl}, of matrix $\mathbf{C}$ for the regular square tessellation superimposed upon a two-dimensional planar surface, which are

$$\lambda_{kl} = 2\left[\cos\left(\frac{k\pi}{P+1}\right) + \cos\left(\frac{l\pi}{Q+1}\right)\right], \quad k = 1, 2, \ldots, P \quad \text{and} \quad l = 1, 2, \ldots, Q, \tag{2}$$

for a $P \times Q$ rectangular geographic region ($n = PQ$).

Gasim [12] shows the derivation of this solution using Kronecker products ($\otimes$), and extends these results to other special adjacency definitions of matrix $\mathbf{C}$.

Griffith and Sone [18] furnish a very good approximation for the Jacobian term $\frac{1}{2}\mathrm{LN}[\det(\mathbf{V})]$, for any planar surface partitioning, that requires the extreme eigenvalues of matrix $\mathbf{C}$ or $\mathbf{W}$, depending upon which matrix is used to define $\mathbf{V}$. In addition, a common measure used in spatial statistics is the Moran coefficient, whose formula involves the matrix $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$. Jong et al. [11] use Rayleigh quotients [20, p. 448] to show that the largest and smallest values that a Moran coefficient can achieve are determined by the extreme eigenvalues of this matrix.

Meanwhile, Tiefelsdorf and Boots [30] derive results for the Moran coefficient based upon the entire set of eigenvalues of matrix $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$, for any planar surface partitioning. Griffith [15,16] extends their findings by showing the spatial statistical importance of the associated set of eigenvectors.

Therefore, spatial statistics furnishes a practical context that motivates the study of properties of matrices $\mathbf{C}$, $\mathbf{W}$, and $\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$.

2. Extension of results for matrices C and W depicting a regular square tessellation

Extending the work of Balisevsky [4, p. 224] from one-dimension to two-dimensions, and following Gasim [12], we get the following:

Theorem 2.1. *Suppose the two-dimensional planar surface partitioning is that of a regular square tessellation forming a $P \times Q$ rectangular geographic region, where P is the number of units in a horizontal axis direction, and Q is the number of units in a vertical axis direction. Let $\mathbf{C}_{PQ}$ denote the binary incidence matrix $\mathbf{C}$, and $\mathbf{E}_{pq}$*

denote the eigenvector matrix $\mathbf{E}$ *for this particular partitioning. Then the eigenvectors of matrix* $\mathbf{C}_{PQ}$ *are given by*

$$\mathbf{E}_{pq} = \left\langle \frac{2}{\sqrt{(P+1)(Q+1)}} \sin\left(\frac{kp\pi}{P+1}\right) \times \sin\left(\frac{lq\pi}{Q+1}\right)\right\rangle,$$
$$k = 1, 2, \ldots, P, \quad l = 1, 2, \ldots, Q, \quad p = 1, 2, \ldots, P, \quad \textit{and} \quad q = 1, 2, \ldots, Q.$$

Proof. Consider the decomposition $\mathbf{C} = \mathbf{C}_{PQ} = (\mathbf{I}_P \otimes \mathbf{C}_Q) + (\mathbf{C}_P \otimes \mathbf{I}_Q)$ [12] and let $\mathbf{\Lambda}_h$ be a diagonal matrix of eigenvalues and $\mathbf{E}_h$ be the accompanying matrix of eigenvectors for matrix $\mathbf{C}_h$. Following Lancaster [22, p. 259], Gasim [12, p. 393] shows that $\mathbf{E}^{\mathrm{T}}\mathbf{C}_{PQ}\mathbf{E} = \mathbf{E}^{\mathrm{T}}[(\mathbf{I}_P \otimes \mathbf{C}_Q) + (\mathbf{C}_P \otimes \mathbf{I}_Q)]\mathbf{E} = \mathbf{I}_P \otimes \mathbf{\Lambda}_Q + \mathbf{\Lambda}_P \otimes \mathbf{I}_Q$.

Now,

$$\begin{aligned}
\mathbf{C}_{PQ} &= (\mathbf{I}_P \otimes \mathbf{C}_Q) + (\mathbf{C}_P \otimes \mathbf{I}_Q) \\
&= (\mathbf{E}_P \mathbf{I}_P {\mathbf{E}_P}^{\mathrm{T}}) \otimes (\mathbf{E}_Q \mathbf{\Lambda}_Q {\mathbf{E}_Q}^{\mathrm{T}}) + (\mathbf{E}_P \mathbf{\Lambda}_P {\mathbf{E}_P}^{\mathrm{T}}) \otimes (\mathbf{E}_Q \mathbf{I}_Q {\mathbf{E}_Q}^{\mathrm{T}}) \\
&= (\mathbf{E}_P \otimes \mathbf{E}_Q)(\mathbf{I}_P \otimes \mathbf{\Lambda}_Q)(\mathbf{E}_P \otimes \mathbf{E}_Q)^{\mathrm{T}} + (\mathbf{E}_P \otimes \mathbf{E}_Q)(\mathbf{\Lambda}_P \otimes \mathbf{I}_Q)(\mathbf{E}_P \otimes \mathbf{E}_Q)^{\mathrm{T}} \\
&= (\mathbf{E}_P \otimes \mathbf{E}_Q)[\mathbf{I}_P \otimes \mathbf{\Lambda}_Q + \mathbf{\Lambda}_P \otimes \mathbf{I}_Q](\mathbf{E}_P \otimes \mathbf{E}_Q)^{\mathrm{T}} \\
&= \mathbf{E}[\mathbf{I}_P \otimes \mathbf{\Lambda}_Q + \mathbf{\Lambda}_P \otimes \mathbf{I}_Q]\mathbf{E}^{\mathrm{T}}
\end{aligned}$$

Therefore, $\mathbf{E} = \mathbf{E}_P \otimes \mathbf{E}_Q$, and hence the eigenvectors of matrix $\mathbf{C}_{PQ}$ are $\mathbf{E}_P \otimes \mathbf{E}_Q$.

But matrix $\mathbf{C}_P$ is a $P \times P$ and $\mathbf{C}_Q$ is a $Q \times Q$ tridiagonal matrix like the ones used in time series analysis; the upper and lower off-diagonals contain ones, and all other cells contain 0's. The eigenvectors of these matrices, respectively, are given by [4, p. 224]

$$\mathbf{E}_p = \left\langle \frac{\sqrt{2}}{\sqrt{P+1}} \sin\left(\frac{kp\pi}{P+1}\right)\right\rangle, \quad k = 1, 2, \ldots, P, \quad p = 1, 2, \ldots, P,$$

and

$$\mathbf{E}_q = \left\langle \frac{\sqrt{2}}{\sqrt{Q+1}} \sin\left(\frac{lq\pi}{Q+1}\right)\right\rangle, \quad l = 1, 2, \ldots, Q, \quad q = 1, 2, \ldots, Q.$$

Therefore,

$$\mathbf{E}_{pq} = \left\langle \frac{2}{\sqrt{(P+1)(Q+1)}} \sin\left(\frac{kp\pi}{P+1}\right) \times \sin\left(\frac{lq\pi}{Q+1}\right)\right\rangle,$$
$$k = 1, 2, \ldots, P, \quad l = 1, 2, \ldots, Q,$$
$$p = 1, 2, \ldots, P, \quad \text{and} \quad q = 1, 2, \ldots, Q. \qquad \square$$

Consequently, all eigenvalues and eigenvectors are known analytically for the binary geographic weights matrix $\mathbf{C}$ affiliated with remotely sensed images, regardless of the size of $n = PQ$.

Martin and Griffith [24] note that if the regular square tessellation is mapped onto a torus, then the eigenvalues given by Eq. (2) are replaced with

$$\lambda_{kl} = 2\left[\cos\left(\frac{k2\pi}{P}\right) + \cos\left(\frac{l2\pi}{Q}\right)\right],$$
$$k = 1, 2, \ldots, P, \quad l = 1, 2, \ldots, Q, \tag{3}$$

for a regular square tessellation partitioning of a $P \times Q$ rectangular geographic region.

Hence, a corollary to Theorem 2.1 is as follows:

Corollary 2.1. *Suppose the two-dimensional surface partitioning is that of a regular square tessellation mapped onto a torus. Then the eigenvectors of matrix* $\mathbf{C}_{PQ}$ *are given by*

$$\mathbf{E}_{pq} = \left\langle \frac{1}{\sqrt{PQ}}\left[\sin\left(\frac{kp2\pi}{P}\right) + \cos\left(\frac{kp2\pi}{P}\right)\right] \times \left[\sin\left(\frac{lq2\pi}{Q}\right) + \cos\left(\frac{lq2\pi}{Q}\right)\right]\right\rangle,$$
$$k = 1, 2, \ldots, P, \quad l = 1, 2, \ldots, Q,$$
$$p = 1, 2, \ldots, P, \quad \text{and} \quad q = 1, 2, \ldots, Q.$$

The proof for this corollary exactly parallels that for Theorem 2.1. The following additional corollary to Theorem 2.1 exploits the well-known eigenfunction property that multiplying a matrix by a constant multiplies that matrix's eigenvalues by the same constant while leaving its eigenvectors unchanged.

Corollary 2.2. *Suppose the two-dimensional surface partitioning is that of a regular square tessellation mapped onto a torus. Then matrix* $\mathbf{W}_{PQ} = \frac{1}{4}\mathbf{C}_{PQ}$, *its eigenvectors are given by Corollary 2.1, and its eigenvalues are*

$$\lambda_{kl} = \left[\cos\left(\frac{k2\pi}{P}\right) + \cos\left(\frac{l2\pi}{Q}\right)\right] \Big/ 2, \quad k = 1, 2, \ldots, P, \quad l = 1, 2, \ldots, Q.$$

These torus results also are a two-dimensional extension of the work of Balisevsky [4, p. 224].

The remaining solution to complete the regular square tessellation case is for matrix **W** depicting a two-dimensional planar surface partitioning. First consider the one-dimensional situation. Let the number of nodes in a linear graph (i.e., a sequence of nodes forming a straight line) be P. Denote the associated matrix **C** by $\mathbf{C}_P$, which is tridiagonal, with 1's on the upper and lower diagonals immediately adjacent to the main diagonal and 0's on the main diagonal itself; this particular matrix also is encountered in time series analyses. Suppose diagonal matrix $\mathbf{D}^{-1}$ has 1 in diagonal cells (1,1) and (P, P), and 0.5 in all of the remaining diagonal cells (i.e., the inverses of the row sums of matrix $\mathbf{C}_P$). The eigenvalues for the resulting stochastic matrix, denoted here as $\mathbf{W}_P = \mathbf{D}^{-1}\mathbf{C}_P$, were first reported widely by Berman and Plemmons [6] and Griffith [14] as

$$\lambda_k = \cos\left(\frac{k\pi}{P-1}\right), \quad k = 0, 1, \ldots, P-1. \tag{4}$$

Hartfiel and Meyer [19, p. 198] also report this result, in a slightly different but insightful way; they separate $k = 0$ and $k = P - 1$ values from the remaining $P - 2$ values. Their derivation inspired the following theorem:

Theorem 2.2. *Suppose $Q = 1$ for a regular square tessellation surface partitioning. Let matrix $\mathbf{W}_P$ denote the stochastic version of matrix $\mathbf{C}_P$, and $\mathbf{E}_k (k = 1, 2, \ldots, P)$ denote the normalized eigenvectors of matrix $\mathbf{W}_P$. Then the eigenvectors of matrix $\mathbf{W}_P$ are given by*

$$\mathbf{E}_1 = \frac{1}{\sqrt{P}}\begin{bmatrix} \cos\left(\frac{0\times 0\times\pi}{P-1}\right) \\ \cos\left(\frac{0\times 1\times\pi}{P-1}\right) \\ \vdots \\ \cos\left[\frac{0\times(P-1)\times\pi}{P-1}\right] \end{bmatrix},$$

$$\mathbf{E}_P = \frac{1}{\sqrt{P}}\begin{bmatrix} \cos(0\times\pi) \\ \cos(1\times\pi) \\ \vdots \\ \cos[(P-1)\times\pi] \end{bmatrix},$$

and

$$\mathbf{E}_k = \sqrt{\frac{2}{P-1}}\begin{bmatrix} \cos\left(\frac{k\times 0\times\pi}{P-1}\right) \\ \cos\left(\frac{k\times 1\times\pi}{P-1}\right) \\ \vdots \\ \cos\left[\frac{k\times(P-1)\times\pi}{P-1}\right] \end{bmatrix}, \quad k = 1, 2, \ldots, P-2.$$

Proof. Solving the standard eigenvector equation $(\mathbf{W}_P - \lambda_k\mathbf{I})\mathbf{E}_k = \mathbf{0}$, the first two equations of the system of n linear equations defined by matrix $\mathbf{W}_P$ are

$$-\cos\left(\frac{k\pi}{P-1}\right)e_{1k} + e_{2k} = 0,$$

$$e_{1k}/2 - \cos\left(\frac{k\pi}{P-1}\right) e_{2k} + e_{3k}/2 = 0,$$

which give

$$\frac{e_{1k}}{1} = \frac{e_{2k}}{\cos\left(\frac{k\pi}{P-1}\right)} = \frac{e_{3k}}{2\cos^2\left(\frac{k\pi}{P-1}\right) - 1}$$

$$\Rightarrow \frac{e_{1k}}{\cos\left(\frac{0k\pi}{P-1}\right)} = \frac{e_{2k}}{\cos\left(\frac{1k\pi}{P-1}\right)} = \frac{e_{3k}}{\cos\left(\frac{2k\pi}{P-1}\right)}.$$

Extending this last result by induction on P proves the vector entries for $\mathbf{E}_k$. Next, from the summation calculus, $\mathbf{\Delta}^{-1}$,

$$\begin{aligned}\sum_{i=0}^{P-1} \cos^2\left(\frac{k\pi i}{P-1}\right) &= 1 + \sum_{i=1}^{P-1} \cos^2\left(\frac{k\pi i}{P-1}\right) \\ &= \frac{1}{2} + \frac{P}{2} - \frac{1}{2}\cos^2\left(\frac{k\pi P}{P-1}\right) \\ &\quad + \frac{1}{2}\frac{\cos\left(\frac{k\pi}{P-1}\right)}{\sin\left(\frac{k\pi}{P-1}\right)} \cos\left(\frac{k\pi P}{P-1}\right) \sin\left(\frac{k\pi P}{P-1}\right) \\ &= \frac{P+1}{2}; \quad k = 1, 2, \ldots, P-2.\end{aligned}$$

The denominator term $\sin(k\pi P/P-1)$ is what restricts k from taking on the values of 0 and $P-1$. In addition, when $k = 0$ all of the vector elements of $\mathbf{E}_1$ are 1, and hence the sum of squares equals P; when $k = P-1$ all of the vector elements of $\mathbf{E}_P$ are ± 1, and again the sum of squares equals P. □

The qualitatively different form of these eigenfunctions corrects a popular misconception appearing in the literature having to do with the asymptotic convergence of $\frac{1}{2}\mathbf{C}_P$ on matrix $\mathbf{W}_P$ (e.g., [12]). Rather, the eigenvalues involve a different cosine argument, whereas the eigenvectors involve the cosine rather than the sine function.

A similarity matrix for $\mathbf{W}_P$ is given by $\mathbf{D}^{-1/2}\mathbf{C}\mathbf{D}^{-1/2}$; in other words, the eigenvalues of both of these matrices are the same. Because the analytical eigenvectors of matrix $\mathbf{W}_P$ are known, then those for matrix $\mathbf{D}^{-1/2}\mathbf{C}\mathbf{D}^{-1/2}$ can be determined.

Theorem 2.3. *Let* $\mathbf{D}$ *be a diagonal matrix whose* (i, i) *entry is* $\sum_{j=1}^{n} c_{ij}$, *where* c_{ij} *is the* (i, j)th *element of matrix* $\mathbf{C}_P$. *Then the eigenvectors of matrix* $\mathbf{D}^{-1/2}\mathbf{C}_P\mathbf{D}^{-1/2}$ *are* $\mathbf{D}^{1/2}\mathbf{E}_k$, *where* $\mathbf{E}_k (k = 1, 2, \ldots, P)$ *are the eigenvectors given by Theorem 2.2.*

Proof. Consider the solution to the eigenvector problem $(\mathbf{W}_P - \lambda_k \mathbf{I})\mathbf{E}_k = \mathbf{0}$ given by Theorem 2.2. Now,

$$(\mathbf{W}_P - \lambda_k \mathbf{I})\mathbf{E}_k = (\mathbf{D}^{-1}\mathbf{C} - \lambda_k \mathbf{I})\mathbf{E}_k = \mathbf{0}$$

$$(\mathbf{D}^{-1/2}\mathbf{C} - \lambda_k \mathbf{D}^{1/2})\mathbf{E}_k = \mathbf{0}$$

$$\Rightarrow\ (\mathbf{D}^{-1/2}\mathbf{C}\mathbf{D}^{-1/2} - \lambda_k)\mathbf{D}^{1/2}\mathbf{E}_k = 0. \qquad \square$$

Of note is that the eigenvectors $\mathbf{D}^{1/2}\mathbf{E}_k$ still need to be normalized.

Following the derivations of Eq. (2) and Theorem 2.1 does not extend the one-dimensional eigenfunction results in Theorems 2.2 and 2.3 to those for their two-dimensional counterpart, though. Using Kronecker product notation, approximate eigenvalues $\hat{\lambda}_{pq}$ for the λ_{pq}'s of matrix $\mathbf{W}_{PQ}$ should be furnished by

$$\hat{\lambda}_{pq} = (\mathbf{1}_P \otimes \boldsymbol{\lambda}_Q + \boldsymbol{\lambda}_P \otimes \mathbf{1}_Q)/2, \tag{5}$$

where $\hat{\lambda}_{pq}$ is the vector of approximated eigenvalues, and λ_P and λ_Q are vectors, whose elements are given by Eq. (4). This approximation is sensible because if $P = 1$ or $Q = 1$, then it is exact, and as $P \to \infty$ and $Q \to \infty$ it converges on the actual eigenvalues. Comparison of Eq. (5) with numerical results yields the following:

Conjecture 1. Consider a $P \times Q$ regular square tessellation planar surface partitioning. Let $\mathbf{W}_{PQ}$ be the stochastic version of matrix $\mathbf{C}_{PQ}$ for a $P \times Q$ regular square tessellation surface partitioning. If $P = Q$, then the eigenvalues given by Eq. (5) for $(p-1)/(P-1) = (q-1)/(Q-1)$ as well as for $(p-1)/(P-1) = (Q-p)/(Q-1)$ are identical to their corresponding eigenvalues extracted from matrix $\mathbf{W}_{PQ}$.

Since the sum of the eigenvalues given by Eq. (4) is 0, then the sum of $\hat{\lambda}_{pq}$ given by Eq. (5) also is 0. In addition, the sum of the squared eigenvalues of matrix $\mathbf{W}_{PQ}$ is given by $\sum_{p=1}^{P}\sum_{q=1}^{Q}\lambda_{pq}^2 = \mathbf{1}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{C}_{PQ}\mathbf{D}^{-1}\mathbf{1}$ [13, Theorem 9.1.14, p. 301]. Therefore, a good approximation for these eigenvalues is given by

$$\begin{aligned}\hat{\lambda}_{pq,\gamma} =& \mathbf{I}_{\text{sign}}[(\mathbf{I} - \mathbf{I}_{\text{correct}})|(\boldsymbol{\lambda}_Q \otimes \mathbf{1}_P + \mathbf{1}_Q \otimes \boldsymbol{\lambda}_P)/2|_{\text{diag}}]^{\gamma}\mathbf{1}_{pq} \\ &+ \mathbf{I}_{\text{correct}}(\boldsymbol{\lambda}_Q \otimes \mathbf{1}_P + \mathbf{1}_Q \otimes \boldsymbol{\lambda}_P)/2,\end{aligned} \tag{6}$$

where $0 < \gamma \leqslant 1$ is selected such that

$$\sum_{p=1}^{P}\sum_{q=1}^{Q}\hat{\lambda}_{pq,\gamma}^2 = \mathbf{1}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{C}_{PQ}\mathbf{D}^{-1}\mathbf{1} = \frac{18PQ + 11P + 11Q + 12}{72},$$

when $P \geqslant 4$ and $Q \geqslant 4$, subscript "diag" denotes a diagonal matrix, "|" denotes absolute value, and the diagonal entries of diagonal matrix $\mathbf{I}_{\text{correct}} = 1$ if an approximate eigenvalue is correct and 0 otherwise, and of diagonal matrix $\mathbf{I}_{\text{sign}} = -1$ if Eq. (5) is negative and 1 otherwise. In addition we have the following:

Conjecture 2. Let $\mathbf{W}_{PQ}$ be the stochastic version of matrix $\mathbf{C}_{PQ}$ for a $P \times Q$ regular square tessellation planar surface partitioning, and let the vectors $\mathbf{E}_p$ and $\mathbf{E}_q$ be those given by Theorem 2.3. Then the eigenvectors of matrix $\mathbf{W}_{PQ}$ are approximately equal to the Kronecker product $\mathbf{E}_p \otimes \mathbf{E}_q (p = 1, 2, \ldots, P; q = 1, 2, \ldots, Q)$.

The results stated in Conjecture 2 at least are asymptotically correct.

2.1. *Means of and correlations between the eigenvectors of matrix* $\mathbf{C}_{PQ}$

Two measures commonly used in statistics are the arithmetic mean and the product moment correlation coefficient. For a set of n variable values denoted by e_i, the mean may be defined as $\sum_{i=1}^{n} e_i/n = \mathbf{1}^{\mathrm{T}}\mathbf{E}/n$; for a set of n paired variable values denoted by e_i and f_i, the product moment correlation coefficient may be written, using matrix notation, as

$$\frac{\mathbf{E}^{\mathrm{T}}\left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{F}}{\sqrt{\mathbf{E}^{\mathrm{T}}\left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{E}}\sqrt{\mathbf{F}^{\mathrm{T}}\left(\mathbf{I} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{F}}}. \tag{7}$$

With regard to the mean, we have the following:

Theorem 2.4. *Suppose a regular square tessellation planar surface partitioning is finite* ($P < \infty$, $Q < \infty$). *Then a number of the eigenvectors*, $\mathbf{E}_{pq}$, *of matrix* $\mathbf{C}_{PQ}$ *have a zero mean.*

Proof. For any eigenvector of matrix $\mathbf{C}_{pq}$,

$$\mathbf{1}^{\mathrm{T}}\mathbf{E}_{pq} = \frac{2}{\sqrt{(P+1)(Q+1)}}\left[\sum_{k=1}^{P}\sin\left(\frac{kp\pi}{P+1}\right)\right] \times \left[\sum_{l=1}^{Q}\sin\left(\frac{lq\pi}{Q+1}\right)\right].$$

Using the summation calculus, $\boldsymbol{\Delta}^{-1}$, for a finite series involving the sine function yields

$$\mathbf{1}^{\mathrm{T}}\mathbf{E}_{pq} = \frac{2}{\sqrt{(P+1)(Q+1)}}\left[-\frac{\cos\left(p\frac{P+\frac{1}{2}}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)} - \frac{-\cos\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}\right]$$

$$\times\left[-\frac{\cos\left(q\frac{Q+\frac{1}{2}}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}-\frac{-\cos\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}\right].$$

Using standard trigonometric identities [29], when p is even, then

$$\left[-\frac{\cos\left(p\frac{P+\frac{1}{2}}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}-\frac{-\cos\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}\right]=0;$$

when p is odd, then

$$\left[-\frac{\cos\left(p\frac{P+\frac{1}{2}}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}-\frac{-\cos\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}{2\sin\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)}\right]=\frac{1}{\tan\left(p\frac{1}{2}\frac{1}{P+1}\pi\right)};$$

when q is even, then

$$\left[-\frac{\cos\left(q\frac{Q+\frac{1}{2}}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}-\frac{-\cos\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}\right]=0;$$

and when q is odd, then

$$\left[-\frac{\cos\left(q\frac{Q+\frac{1}{2}}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}-\frac{-\cos\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}{2\sin\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}\right]=\frac{1}{\tan\left(q\frac{1}{2}\frac{1}{Q+1}\pi\right)}.$$

Therefore, when either p or q is even, then the affiliated vector sum is $\mathbf{1}^{\mathrm{T}}\mathbf{E}_{pq} \equiv 0$. □

When both P and Q are even then the number of eigenvectors whose sums are exactly zero is $3PQ/4$. When both P and Q are odd, this number is $(3PQ-P-Q-1)/4$. When P is even and Q is odd, this number is $(P(3PQ-1))/4$; and when P is odd and Q is even, this number is $(Q(3PQ-1))/4$.

With regard to the correlation coefficient, we have the following:

Theorem 2.5. *Let $n = PQ$, the number of units into which a surface is partitioned. Then as both P and Q go to infinity, all of the correlations (ρ) amongst the nonprincipal eigenvectors of matrix $\mathbf{C}_{PQ}$ converge on zero.*

Proof. Substituting two distinct nonprincipal eigenvectors of matrix $\mathbf{C}_{PQ}$, say $\mathbf{E}_j$ and $\mathbf{E}_k (j \neq k)$, into expression (6) yields

$$\frac{\mathbf{E}_j^{\mathrm{T}}\left(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{E}_k}{\sqrt{\mathbf{E}_j^{\mathrm{T}}\left(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{E}_j}\sqrt{\mathbf{E}_k^{\mathrm{T}}\left(\mathbf{I}-\frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{n}\right)\mathbf{E}_k}}$$

$$=-\frac{\mathbf{E}_j^{\mathrm{T}}\mathbf{1}\left(\frac{\mathbf{1}^{\mathrm{T}}\mathbf{E}_k}{n}\right)}{\sqrt{1-\mathbf{E}_j^{\mathrm{T}}\mathbf{1}\left(\frac{\mathbf{1}^{\mathrm{T}}\mathbf{E}_{jk}}{n}\right)}\sqrt{1-\mathbf{E}_k^{\mathrm{T}}\mathbf{1}\left(\frac{\mathbf{1}^{\mathrm{T}}\mathbf{E}_k}{n}\right)}}.$$

Theorem 2.4 states that for any $j = pq$ or $k = rs$ having p, q, r or s even, the numerator of this expession is 0 regardless of the magnitude of P and Q, and hence the correlation coefficient is 0. If p, q, r and s all are odd, the numerator contains the product of four terms of the form

$$\frac{\frac{1}{T}\times\frac{2}{T+1}}{\tan\left[\frac{t\pi}{2(T+1)}\right]}.$$

As $t \to T$ and $T \to \infty$, the denominator of this term goes to ∞. As $t \to T/2$ and $T \to \infty$, the denominator of this term goes to 1, but the numerator goes to 0. As $t \to 1$, by L‘Hospital’s rule (from calculus) the limit of this expression is equivalent to

$$\lim_{T\to\infty}\frac{-2\frac{2T+1}{T^2(T+1)^2}}{-\frac{1}{2}\frac{t\pi}{(T+1)^2\cos^2\left[\frac{t\pi}{2(T+1)}\right]}},$$

which is 0.

Therefore, since all of the means of the nonprincipal eigenvectors asymptotically go to zero, these nonprincipal eigenvectors asymptotically are mutually uncorrelated. □

Of note is that these asymptotic eigenvectors are both orthogonal and uncorrelated.

Another noteworthy correlation is revealed by the following theorem:

Theorem 2.6. *Let* $\mathbf{E}_n$ *and* $\mathbf{E}_n{}^*$, *respectively, denote the normalized eigenvectors associated with the smallest eigenvalues of matrices* $\mathbf{C}$ *and* $(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)$. *By Theorem* 3.1, *vector* $\mathbf{E}_n{}^*$ *is approximated by* $k(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_n$. *Then vector* $\mathbf{E}_n{}^*$ *has a correlation* (ρ) *between itself and* $\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_n{}^*$ *of exactly* -1.

Proof. Consider the regression equation $\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_n{}^* = \boldsymbol{\alpha}\mathbf{1}+\boldsymbol{\beta}\mathbf{E}_n{}^*$.

$$(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_n{}^* = \boldsymbol{\beta}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_n{}^*.$$

From ordinary least-squares (OLS) regression theory [25],

$$\hat{\boldsymbol{\beta}} = [\mathbf{E}_n{}^{*\mathrm{T}}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}_n{}^{*}]^{-1}\mathbf{E}_n{}^{*\mathrm{T}}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}_n{}^{*}$$
$$= \mathbf{E}_n{}^{*\mathrm{T}}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}_n{}^{*} = \lambda_n.$$

Hence expression $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$ is a regression coefficient, which can be rewritten as

$$\hat{\boldsymbol{\beta}} = \rho \frac{\mathbf{S}_{\mathbf{C}\left(\mathbf{I} - \frac{\mathbf{11}^{\mathrm{T}}}{n}\right)\mathbf{E}*}}{\mathbf{S}_{\left(\mathbf{I} - \frac{\mathbf{11}^{\mathrm{T}}}{n}\right)\mathbf{E}*}} = \boldsymbol{\rho}|\lambda_n|.$$

Therefore, since both $\mathbf{E}_n{}^{*}$ and $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}_n{}^{*}$ are centered, and $\mathbf{E}_n{}^{*}$ is normalized, then

$$\text{Min}: \quad \mathbf{E}_n{}^{*T}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}_n{}^{*} \Rightarrow \rho = -1. \qquad \square$$

Of course the largest eigenvalue of matrix **C** as well as its associated eigenvector can be quickly calculated, even for very large *n*, using one of the oldest and the well-known method of

$$\lim_{k\to\infty} \frac{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{k+1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{k}\mathbf{1}} = \lambda_1$$

[7, p. 213]. Theorem 2.6 furnishes analytical information that aids in calculating the minimum eigenfunction of matrix **C**.

3. Commonalities between eigenfunctions of matrices C and $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n) \times \mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$

Mäkeläinen [23] identifies a number of commonalities between matrices of this sort, but for when matrix **C** is either positive or nonnegative-definite. Neither matrix **C** nor matrix **W** is positive or nonnegative-definite. Hence Mäkeläinen's results provide some guidance here (e.g., a suitable diagonal matrix a**I** can be added to either matrix **C** or **W** to make it positive-definite, without altering the associated eigenvectors), but do not necessarily directly apply. Perhaps the most relevant result by Mäkeläinen is his Theorem 4.2 [23, p. 34].

The following theorem is posited as a description of the relationship between the eigenvectors of matrix **C** and those of expression $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$.

Theorem 3.1. *Suppose that symmetric binary* (0–1) *matrix* **C** *is an* $n \times n$ *incidence matrix representing a planar partitioning of some two-dimensional geographic surface into n polygonal units. Let* **E** *be the matrix of eigenvectors of* $\mathbf{C}$, $\mathbf{E}_1$ *denote the principal eigenvector in matrix* $\mathbf{E}$, $\mathbf{I}$ *be the identity matrix*, $\mathbf{1}$ *be an* $n \times 1$ *vector of ones, and k be a positive constant. If the nonprincipal eigenvectors of matrix* **C** *are centered and renormalized* (*i.e.*, *replaced with* $k(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{E}$), *then as n increases they converge upon the eigenvectors of matrix* $(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I} - \mathbf{11}^{\mathrm{T}}/n)$, *but with the centered principal eigenvector* $(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_1$ *being replaced with vector* $1/\sqrt{n}\mathbf{1}$.

Proof. Consider the eigenvector problem $(\mathbf{C}-\lambda\mathbf{I})\mathbf{E}=\mathbf{0}$, s.t. $\mathbf{E}^{\mathrm{T}}\mathbf{E}=\mathbf{I}$.

$$
\begin{aligned}
&(\mathbf{C}-\lambda\mathbf{I})\mathbf{E}-(\mathbf{C}-\lambda\mathbf{I})(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}=-(\mathbf{C}-\lambda\mathbf{I})(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E},\\
&(\mathbf{C}-\lambda\mathbf{I})(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}=-(\mathbf{C}-\lambda\mathbf{I})(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E},\\
&(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)[\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)-\lambda(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)](\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}\\
&\quad=-(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)(\mathbf{C}-\lambda\mathbf{I})(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E},\\
&[(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)-\lambda\mathbf{I}](\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}+\lambda(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}\\
&\quad=-(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)(\mathbf{C}-\lambda\mathbf{I})(\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E},\\
&[(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)-\lambda\mathbf{I}](\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}\\
&\quad=-(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}\mathbf{1}(\mathbf{1}^{\mathrm{T}}\mathbf{E}/n).
\end{aligned}
$$

Since $\mathbf{C1}$ is a vector of finite-entries, and $\mathbf{E}^{\mathrm{T}}\mathbf{E}=\mathbf{I}$, then $\mathbf{1}^{\mathrm{T}}\mathbf{E}_k\leqslant\sqrt{n}$, and hence

$$\lim_{n\to\infty}\mathbf{1}^{\mathrm{T}}\mathbf{E}/n=0.$$

Thus

$$[(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)-\lambda\mathbf{I}](\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}_k\to 0,\ k\neq 1.$$

If $\mathbf{E}_1\geqslant\mathbf{0}$ (guaranteed by the Perron–Frobenius theorem) is replaced with $1/\sqrt{n}\mathbf{1}$, then $[(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)-\lambda\mathbf{I}](\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\left(1/\sqrt{n}\mathbf{1}\right)=\mathbf{0}$, with $\lambda=0$

Hence, asymptotically $k(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{E}$ are the eigenvectors of $(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\times\mathbf{C}(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)$ once $\mathbf{E}_1$ is replaced with $1/\sqrt{n}\mathbf{1}$. □

Accordingly, all but the largest eigenvalue of matrix $\mathbf{C}$ approximately equal and asymptotically converge upon one of the eigenvalues of matrix $(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)\mathbf{C}\times(\mathbf{I}-\mathbf{1}\mathbf{1}^{\mathrm{T}}/n)$.

4. Approximate eigenvalues associated with a $P\times Q$ regular hexagonal tessellation

Techniques used to obtain results for the regular square tessellation case provide guidance for the regular hexagonal tessellation situation. Here the variance term for the eigenvalues extracted from matrix $\mathbf{C}_{PQ}$ based upon a planar surface may be calculated with

$$\mathbf{1}^{\mathrm{T}}\mathbf{C}\mathbf{1}=2(3PQ-2P-2Q+1).\tag{8}$$

In addition, the limit of the principal eigenvalue, λ_1, is 6 while the limit of the minimum eigenvalue, λ_n is -3. With the eigenvalues of a square tessellation given by the well-known equation

$$\lambda_{pq} = 2\left[\cos\left(\frac{p\pi}{P+1}\right) + \cos\left(\frac{q\pi}{Q+1}\right)\right],$$

and those for the queen's definition (another analogy with chess moves) of adjacency given by

$$\lambda_{pq} = 2\left[\cos\left(\frac{p\pi}{P+1}\right) + \cos\left(\frac{q\pi}{Q+1}\right) + 2\cos\left(\frac{p\pi}{P+1}\right)\times\cos\left(\frac{q\pi}{Q+1}\right)\right],$$

following the type of argument outlined by Gasim [12], those for a hexagonal tessellation should be approximated by

$$\hat{\lambda}_{pq} = 2\cos\left(\frac{p\pi}{P+1}\right) + 2\cos\left(\frac{q\pi}{Q+1}\right) + 2\cos\left(\frac{p\pi}{P+1}\right)\times\cos\left(\frac{q\pi}{Q+1}\right). \tag{9}$$

Eq. (9) supplies the basis for a good approximation of the actual eigenvalues.

The extreme eigenvalues of a regular hexagonal tessellation can be accurately estimated, being constrained to achieve their respective asymptotic values, with

$$\hat{\lambda}_{\max} = 0.850393\left\{2\left[\cos\left(\frac{\pi}{P+1}\right) + \cos\left(\frac{\pi}{Q+1}\right)\right]\right\} + 0.3248037 \times\left\{2\left[\cos\left(\frac{\pi}{P+1}\right) + \cos\left(\frac{\pi}{Q+1}\right) + 2\cos\left(\frac{\pi}{P+1}\right)\times\cos\left(\frac{\pi}{Q+1}\right)\right]\right\}, \tag{10}$$

and

$$\hat{\lambda}_{\min} = 0.30959 + 0.82713\left\{2\left[\cos\left(\frac{P\pi}{P+1}\right) + \cos\left(\frac{Q\pi}{Q+1}\right)\right]\right\}, \tag{11}$$

whose coefficients have been estimated using nonlinear regression techniques. The limit of Eq. (10) is 6, and for a systematic sample of hexagonal surface partitionings from $P = 5$ and $Q = 5$, to $P = 50$ and $Q = 50$, both the description of the maximum eigenvalue is good (multiple correlation $R^2 = 0.991$) and the residuals of the estimation equation are statistically well-behaved. Eq. (10) performs better on larger P and Q values, which in fact are the ones of most concern, since the eigenvalues of matrix $\mathbf{C}$ for relatively small P and Q values can be easily calculated numerically. Meanwhile, the limit of Eq. (11) is -2.99894, which is very close to -3. Again, both the description of the minimum eigenvalue is good (multiple correlation $R^2 = 0.981$) and the residuals of the estimation equation are statistically well-behaved.

Now the unknown λ_{pq}'s can be approximated iteratively as follows:

Step 1: Compute $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$, respectively, using Eqs. (10) and (11);

Step 2: Compute the $n = PQ$ estimates $\hat{\lambda}^*_{pq}$ using the following equation (based upon (9)):

$$2\cos\left(\frac{p\pi}{P+1}\right) + \frac{5}{2}\cos\left(\frac{q\pi}{Q+1}\right) + \cos\left(\frac{36.5^\circ - \frac{6.6}{P+1} - \frac{2(6.6)}{Q+1}}{180^\circ}\pi\right) 2\cos\left(\frac{p\pi}{P+1}\right) \times \cos\left(\frac{q\pi}{Q+1}\right),$$

and then determine $\max\{\hat{\lambda}^*_{pq}\}$ and $\min\{\hat{\lambda}^*_{pq}\}$; and

Step 3: Iteratively choose a $\hat{\gamma}$ such that the sum of the following eigenvalue estimates equals 0:

$$\hat{\lambda}_{pq,\gamma} = \left[\frac{\hat{\lambda}^*_{pq} + |\min(\hat{\lambda}^*_{pq})|}{\max(\hat{\lambda}^*_{pq}) + |\min(\hat{\lambda}^*_{pq})|}\right]^{\hat{\gamma}} (\hat{\lambda}_{\max} + |\hat{\lambda}_{\min}|) + \hat{\lambda}_{\min}.$$

Eq. (8) furnishes a check on these estimates, for good ones will have a sum of squares close to the value yielded by Eq. (8).

4.1. Approximate eigenvalues of matrix **W** *for a $P \times Q$ regular hexagonal tessellation*

To begin, the principal eigenvalue, λ_1, is known theoretically to equal 1. Meanwhile, the minimum eigenvalue may be estimated with the following equation:

$$\hat{\lambda}_{\min} = -0.573089 - \left[\frac{44.81491}{(P+1.46199)^{4.96963}} I_{P,\text{even}} + \frac{45.54852}{(P+3.00664)^{4.16429}} I_{P,\text{odd}} + \frac{44.81491}{(Q+1.27039)^{5.20842}} I_{Q,\text{even}} + \frac{45.54852}{(Q+10.01437)^{2.90953}} I_{Q,\text{odd}}\right] \Big/ 2, \tag{12}$$

whose coefficients have been estimated using nonlinear regression techniques, and which has a corresponding R^2 value of 0.998, but not well-behaved residuals. Of note here is that convergence of this estimate in the limit tracks one trajectory for even values of P and Q, and a second trajectory for odd values of P and Q. It suggests the following conjecture:

Conjecture 3. Let $\mathbf{W}_{PQ}$ be the stochastic matrix version of the binary incidence matrix $\mathbf{C}_{PQ}$ representing a $P \times Q$ regular hexagonal tessellation planar surface partitioning. Then the minimum eigenvalue of matrix $\mathbf{W}$ asymptotically is approximately -0.573089.

Following the developments outlined in the preceding section for estimating the eigenvalues of matrix $\mathbf{C}_{PQ}$, $\hat{\lambda}^*_{pq}$ initially can be estimated with the expression

$$0.35\cos\left[\frac{(p-1)\pi}{P-1}\right] + 0.35\cos\left[\frac{(q-1)\pi}{Q-1}\right] + 0.30\cos\left[\frac{(p-1)\pi}{P-1}\right] \times \cos\left[\frac{(q-1)\pi}{Q-1}\right], \tag{13}$$

and then iteratively improved by estimating $\hat{\gamma}$ such that the sum of

$$\hat{\lambda}_{pq,\gamma} = \left[\frac{\hat{\lambda}^*_{pq} + |\min(\hat{\lambda}^*_{pq})|}{\max(\hat{\lambda}^*_{pq}) + |\min(\hat{\lambda}^*_{pq})|}\right]^{\hat{\gamma}} (1 + |\hat{\lambda}_{\min}|) + \hat{\lambda}_{\min}$$

is 0, where $\hat{\lambda}_{\min}$ is given by Eq. (12). Again, good approximations will have a sum of squares close to

$$\mathbf{1}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{C}_{PQ}\mathbf{D}^{-1}\mathbf{1} = \frac{30PQ + 25P + 24Q + 23}{180},$$

where the horizontal axis is $P \geqslant 4$, the vertical axis is $Q \geqslant 4$, and one pair of parallel hexagon boundaries is orthogonal to the horizontal axis.

5. Conclusion

In conclusion, the spatial statistics literature is replete with linear algebra and its applications. This paper contributes eigenfunction results that are useful to that segment of the spatial statistics literature. It focuses on two versions of incidence matrices commonly employed in geographic analysis that can be viewed in a graph theoretic context, especially in terms of planar graphs. The study of eigenfunctions of graphs has a long history. Hence, the research summarized in this paper also contributes to this graph theory history. With particular reference to linear algebra, results summarized in this paper allow the determinant of selected massively large $n \times n$ matrices to be accurately approximated; these results are relevant to two spatial statistical problem areas, namely remotely sensed data for which the stochastic matrix $\mathbf{W}$ is to be used and regular hexagonal tessellation data such as that for EMAP. In addition, the behavior of eigenfunctions for graphs affiliated with a linear configuration of connected nodes are better understood, as are those for the graphs associated with regular square and hexagonal tessellations.

References

[1] L. Anselin, Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht, 1988.
[2] N. Ahuja, B. Schachter, Pattern Models, Wiley, New York, 1983.
[3] Z. Bai, M. Fahey, G. Golub, Some large-scale matrix computation problems, J. Comput. Appl. Math. 74 (1996) 71–89.
[4] A. Basilevsky, Applied Matrix Algebra in the Statistical Sciences, North-Holland, New York, 1983.
[5] R. Barry, R. Pace, Monte Carlo estimates of the log determinant of large sparse matrices, Linear Algebra Appl. 289 (1999) 41–54.
[6] A. Berman, R. Plemmons, Nonnegative Matrices in the Mathematical Sciences (corrected republication of the 1979 edition), SIAM, Philadelphia, PA, 1994.
[7] F. Chatelin, Eigenvalues of Matrices (translated by W. Ledermann), Wiley, New York, 1993.
[8] F. Chung, Spectral Graph Theory, AMS Providence, RI, 1997.
[9] N. Cressie, Statistics for Spatial Data, Wiley, New York, 1991.
[10] N. Cressie, A. Olsen, D. Cook, Massive data sets: problems and possibilities with application to environmental monitoring, in: Massive Data Sets, edited by Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences, National Research Council, National Academy Press, Washington DC, 1996, pp. 115–119.
[11] P. de Jong, C. Sprenger, F. van Veen, On extreme values of Moran's I and Geary's c, Geograph. Anal. 16 (1984) 17–24.
[12] A. Gasim, First-order autoregressive models: a method for obtaining eigenvalues for weighting matrices, J. Statis. Plan. Infer. 18 (1988) 391–398.
[13] F. Graybill, Matrices with Applications in Statistics, second ed. Wadsworth, Belmont, CA, 1969.
[14] D. Griffith, Towards a theory of spatial statistics, Geograph. Anal. 12 (1980) 325–339.
[15] D. Griffith, Spatial autocorrelation and eigenfunctions of the geographic weights matrix accompanying geo-referenced data, The Canadian Geographer 40 (1996) 351–367.
[16] D. Griffith, A linear regression solution to the spatial autocorrelation problem, J. Geographi. Syst., 2000 (forthcoming).
[17] D. Griffith, L. Layne, A Casebook for Spatial Statistical Data Analysis, Oxford University Press, New York, 1999.
[18] D. Griffith, A. Sone, Trade-offs associated with normalizing constant computational simplifications for estimating spatial statistical models, J. Stat. Comput. Simulat. 51 (1995) 165–183.
[19] D. Hartfiel, C. Meyer, On the structure of stochastic matrices with a subdominant eigenvalue near 1, Linear Algebra Appl. 272 (1998) 193–203.
[20] J. Jackson, A User's Guide to Principal Components, Wiley, New York, 1991.
[21] R. Johnson, D. Wichern, Applied Multivariate Statistical Analysis, third ed., Prentice-Hall, Englewood Cliffs, NJ, 1992.
[22] P. Lancaster, Theory of Matrices, Academic Press, New York, 1977.
[23] T. Mäkeläinen, Extrema for charasteristic roots of product matrices, Commentationes Physico-Mathematicae 38 (4) (1970) 27–53.
[24] R. Martin, D. Griffith, Some results on graph spectra, with applications to geographical spatial modelling, Research Report No. 461/95, Department of Probability and Statistics, University of Sheffield, 1995.
[25] D. Montgomery, E. Peck, Introduction to Linear Regression Analysis, Wiley, New York, 1982.
[26] J. Messer, R. Linhurst, W. Overton, An EPA program for monitoring ecological status and trends, Environmental Monitoring and Assessment 17 (1991) 67–78.

[27] S. Openshaw, Some ideas about the exploratory spatial analysis technology required for massive databases, in: Massive Data Sets edited by the Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences, National Research Council, National Academy Press, Washington, DC, 1996, pp. 149–156.
[28] J. Ord, Estimation methods for models of spatial interaction, J. Amer. Stat. Assoc. 70 (1975) 120–126.
[29] M. Spiegel, Mathematical Handbook of Formulas and Tables, McGraw-Hill, New York, 1968.
[30] M. Tiefelsdorf, B. Boots, The exact distribution of Moran's I, Environment and Planning A 27 (1995) 985–999.