Spatial design matrices and associated quadratic forms: structure and properties

Hillier, Grant and Martellosio, Federico

University of Southampton

2006

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Multivariate Analysis 97 (2006) 1 – 18

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Spatial design matrices and associated quadratic forms: structure and properties

Grant Hillier*, Federico Martellosio

Division of Economics, School of Social Sciences, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

Received 1 May 2003
Available online 7 January 2005

Abstract

The paper provides significant simplifications and extensions of results obtained by Gorsich, Genton, and Strang (J. Multivariate Anal. 80 (2002) 138) on the structure of spatial design matrices. These are the matrices implicitly defined by quadratic forms that arise naturally in modelling intrinsically stationary and isotropic spatial processes. We give concise structural formulae for these matrices, and simple generating functions for them. The generating functions provide formulae for the cumulants of the quadratic forms of interest when the process is Gaussian, second-order stationary and isotropic. We use these to study the statistical properties of the associated quadratic forms, in particular those of the classical variogram estimator, under several assumptions about the actual variogram.

AMS 1991 subject classification: Primary: 62H11; Secondary: 62H10

Keywords: Cumulant; Intrinsically stationary process; Kronecker product; Quadratic form; Spatial design matrix; Variogram

1. Introduction

In modelling spatial data—in general in d dimensions—observed at sites labelled by points in some subset of R^d, it is often assumed that the process is intrinsically stationary and isotropic (see below and [6]). Such models are then—intuitively at least—generalizations

* Corresponding author.
E-mail addresses: ghh@soton.ac.uk (G. Hillier), fm200@soton.ac.uk (F. Martellosio).

doi:10.1016/j.jmva.2004.11.004

of familiar stationary time series models defined on the line (the case $d = 1$), and, we shall see that there is quite a formal structure that reflects this relationship (Theorem 1 below).

In this paper, as in the recent paper by Gorsich et al. [10] (hereafter abbreviated to GGS), we assume that the observational sites are located on a uniform grid in R^d, with n sites on each of d axes. Sites may then be labelled by elements of the set $\Gamma = \Gamma(n, d)$ of sequences $\alpha = (\alpha(1), \ldots, \alpha(d))$ of non-negative integers satisfying $0 \leqslant \alpha(i) \leqslant (n-1)$ for $i = 1, \ldots, d$, and, to avoid ambiguity, we order the sequences in Γ *lexicographically*. Extensions to the case of a rectangular grid are straightforward, but for simplicity we confine our results to the hypercubic grid.

Denoting the observed process by $\{Z(\alpha); \alpha \in \Gamma\}$, intrinsic stationarity entails the assumptions that $E(Z(\alpha))$ is constant, and that, for $\alpha \neq \beta, \gamma(\alpha, \beta) = Var(Z(\alpha) - Z(\beta))$ depends on (α, β) only through $(\alpha - \beta)$, and the isotropy assumption that $\gamma(\alpha, \beta)$ depends on (α, β) only through $h = \|\alpha - \beta\|^2$, the squared Euclidean distance between the sites α and β. In that case the function $2\gamma(h)$ defined by

$$2\gamma(h) = Var(Z(\alpha) - Z(\beta)) \tag{1}$$

is called the *variogram* of the process $Z(\alpha)$. Note that, here and throughout, we use h to denote the *squared* Euclidean distance $\|\alpha - \beta\|^2 = \sum_{i=1}^{d}(\alpha(i) - \beta(i))^2$ between sites, rather than (as is more common) $\|\alpha - \beta\|$ itself. This is notationally more convenient later. Henceforth we take h to be strictly positive unless otherwise indicated.

The natural estimator for $2\gamma(h)$ is based on the function

$$q_h = \sum_{N(h)} (z(\alpha) - z(\beta))^2, \tag{2}$$

where $z(\alpha)$ denotes the observed value of $Z(\alpha)$, and $N(h)$ is the set of (unordered) pairs (α, β) satisfying $\|\alpha - \beta\|^2 = h$. Note that both $\gamma(0) = 0$ and $q_0 = 0$. Statistics of this form are also of interest more generally in the context of modelling spatial processes.

For $h > 0$ the expression on the right in (2) may be written as a quadratic form

$$q_h = z' L_h z = z'(D_h - A_h)z, \tag{3}$$

where $z = (z(\alpha); \alpha \in \Gamma)$ denotes the N-dimensional vector of observations, L_h and A_h are symmetric, and D_h is a diagonal matrix. Here and throughout $N = n^d = |\Gamma|$, the cardinality of Γ, denotes the total sample size. The matrix of this quadratic form, L_h, is the $N \times N$ *spatial design matrix* at distance $\sqrt{h}$, and D_h and $-A_h$ are, respectively, the diagonal and off-diagonal parts of L_h. By expanding the right side of (2) it is easy to see that A_h has a one in positions labelled by pairs (α, β) satisfying $\|\alpha - \beta\|^2 = h$, and zeros elsewhere, and that the diagonal element in row α of D_h is the number of sequences $\beta \in \Gamma$ satisfying $\|\alpha - \beta\|^2 = h$, i.e., the sum of the elements in row α of A_h. The matrices $L_h = L_h(n, d)$ in (3) are, in GGS, denoted by $A^{(d)}(n^d, h)$, with $h = \|\alpha - \beta\|$. The matrix A_h may be interpreted as the adjacency matrix of a graph $G(\Gamma, h)$ with vertex set Γ and edges the pairs $(\alpha, \beta) \in \Gamma \times \Gamma$ for which $\|\alpha - \beta\|^2 = h$. In that context L_h is known as the Laplacian matrix of the graph $G(\Gamma, h)$ (see [17], for instance). Statistics of the type (2) have been studied extensively for the case $d = 1$, beginning with von Neumann et al. [19].

As already mentioned, an important application of the quadratic forms q_h is to the estimation of the variogram in geostatistics. Let $N_h = |N(h)|$ denote the cardinality of the set $N(h)$. The statistic $2\hat{\gamma}(h) = q_h/N_h$, is an unbiased estimator of $2\gamma(h)$, and is often referred to as the classical variogram estimator (see Section 3.2 below, and GGS and the references therein). However, for other purposes it is also of interest to consider the statistics

$$q_h^* = 2 \sum_{N(h)} z(\alpha)z(\beta) = z' A_h z, \tag{4}$$

based on just the off-diagonal part of L_h. To give just a few examples: (i) the statistic q_h^*, normalized by $z'z$, is used to test for spatial autocorrelation at distance $\sqrt{h}$ (see [18]); (ii) if the covariance matrix of the process belongs to the linear span of (some of) the matrices A_h, that is, if the spatial process is not only intrinsically stationary and isotropic, but also second-order stationary, the statistic $q_h^*/(2N_h)$ is (when the process has zero mean) an unbiased estimator of the covariance function at distance $\sqrt{h}$ (see Section 3.2); (iii) if the process is assumed to be Gaussian with precision matrix (inverse covariance matrix) that is a linear combination of matrices I_N and $\{A_h, h \in H_p\}$, where H_p contains p distinct values of h and I_N denotes the $N \times N$ identity matrix, then a pth order conditional autoregression is obtained [4]. The matrices $A_h, h \in H_p$, play the role of spatial weights matrices, and the quadratic forms $(z'z, q_h^*,\ h \in H_p)$, are minimal sufficient statistics for the parameters of the model, and thus form the basis for inference on those parameters.

The problem of interest here is to give structural formulae for the matrices A_h, and thereby for D_h and L_h. Thus, we continue the work of GGS, whose aim was to analyze the eigenstructure of the matrices L_h, with a view to deducing the properties of statistics like q_h and q_h^*, or more specifically of the variogram estimator $2\hat{\gamma}(h)$. It is well-known that under Gaussian assumptions (and also more generally) the properties of q_h and q_h^* depend upon L_h and A_h, respectively, only through their eigenvalues. Our purpose in the present paper will be to simplify and extend the results given in GGS.

In Section 2, we first provide a complete structural representation of the matrices A_h and L_h, and then give generating functions that make their computation straightforward with a standard symbolic computation package. In principle this completely solves the eigenvalue problem, but in practice, since N is usually quite large, direct computation of the eigenvalues would be unreliable. And, as we shall see, except in special cases, both A_h and L_h are sums of non-commuting matrices. Since, in this case, it is generally not possible to express the eigenvalues of the sum in terms of those of the summands, general explicit formulae for the eigenvalues are unlikely to be accessible.

Fortunately, our generating function results do permit the computation of the cumulants of the statistics of interest very simply and directly. In Section 3, we use these expressions to study the properties of the statistics q_h and q_h^* under the assumption that the process $\{Z(\alpha), \alpha \in \Gamma\}$ is Gaussian, second-order stationary, and isotropic. In particular, in Section 3.3 we show that the earlier results can be applied to the study of the properties of the classical variogram estimator $2\hat{\gamma}(h)$ under a variety of assumptions on the actual variogram $2\gamma(h)$.

2. The matrices A_h, D_h and L_h

In this section we give the main structural results for the matrices A_h, D_h and L_h. The elements of these matrices, indexed by pairs $(\alpha, \beta) \in \Gamma \times \Gamma$, are completely determined by n, d and h. The results express these matrices in $d > 1$ dimensions in terms of sums of Kronecker products of the corresponding matrices in dimension $d = 1$. We begin with the key result—a very simple structural formula for the matrices A_h.

2.1. Off-diagonal part

The matrices A_h are defined by

$$(A_h)_{\alpha,\beta} = \begin{cases} 1 \text{ if } \parallel \alpha - \beta \parallel^2 = h, \\ 0 \text{ otherwise.} \end{cases} \tag{5}$$

Evidently, setting $A_0 = I_N$, $\sum_{h \geqslant 0} A_h = J_N$, where J_q is the $q \times q$ matrix with all elements one. In dimension $d = 1$ we denote the $n \times n$ matrices A_{r^2} by F_r, $r = 0, 1, \ldots, n-1$. That is,

$$(F_r)_{i,j} = \begin{cases} 1 \text{ if } |i-j| = r, \\ 0 \text{ otherwise.} \end{cases} \tag{6}$$

Since $\sum_{r=0}^{n-1} F_r = J_n$, we have that

$$J_N = \bigotimes_1^d J_n = \bigotimes_1^d \left(\sum_{r=0}^{n-1} F_r \right) = \sum_{\alpha \in \Gamma} (F_{\alpha(1)} \otimes F_{\alpha(2)} \otimes \cdots \otimes F_{\alpha(d)}) \tag{7}$$

by the multilinearity of the Kronecker (or direct) product '$\otimes$'. Note that the elements of

$$F_\alpha^{\otimes} = F_{\alpha(1)} \otimes F_{\alpha(2)} \otimes \cdots \otimes F_{\alpha(d)} \tag{8}$$

are zeros and ones, so *exactly one term* $F_\alpha^{\otimes}$ on the right in (7) has a one in any given position (β, δ). In view of (7), the following result is not surprising:

Proposition 1. *Let* $\Gamma_h = \{\alpha \in \Gamma : \|\alpha\|^2 = h\}$. *Then*:

$$A_h = \sum_{\alpha \in \Gamma_h} F_\alpha^{\otimes}. \tag{9}$$

Proof. For each pair $(\beta, \delta) \in \Gamma \times \Gamma$, define $\alpha \in \Gamma$ by $\alpha(i) = |\beta(i) - \delta(i)|,\ i = 1, \ldots, d$. From the definition of A_h, $(A_h)_{\beta,\delta} = 1$ if and only if $\|\alpha\|^2 = h$, or $\alpha \in \Gamma_h$. On the other hand, the (β, δ) element of $(F_{\alpha(1)} \otimes F_{\alpha(2)} \otimes \cdots \otimes F_{\alpha(d)})$ is one if and only if

$$|\beta(i) - \delta(i)| = \alpha(i), \ \ for\ i = 1, \ldots, d. \tag{10}$$

Summing the $F_\alpha^{\otimes}$ over Γ_h must therefore yield A_h by the remark following (8). $\square$

For example, if $h = 1$, Γ_1 consists of d sequences containing a single one and $d - 1$ zeros, so that

$$A_1 = \sum_{i=1}^{d} (I_n \otimes \cdots \otimes F_1 \otimes \cdots \otimes I_n)$$

with F_1 in the ith position in the ith term (see the discussion of Eq. (9) in GGS). Likewise, for $h = 2$, Γ_2 consists of the $\binom{d}{2}$ sequences that contain 2 ones and $d - 2$ zeros, so in the corresponding expression for A_2 each term in the sum contains F_1 twice. Notice that, in both of these low-order cases, all the sequences that appear in Γ_h are permutations of a single sequence.

An alternative proof of Proposition 1 based on known graph-theoretic results is worth recording, because it shows immediately how to generalize the result to cover index sets more complex than the uniform grid Γ, e.g., the rectangular grid mentioned in the Introduction. We refer the reader to Cvetković et al. [7] for more on the graph-theoretic details.

Given graphs $G_i(V_i, E_i)$, $i = 1, \ldots, d$, with vertex sets V_i and edge sets E_i, the *direct product* of the G_i, $G_1 \times \cdots \times G_d$ is the graph $G_d^{\times}$, say, defined as follows. The vertex set of $G_d^{\times}$ is the Cartesian product $V_d^{\times} = V_1 \times \cdots \times V_d$ of the V_i, and if $x_i, y_i \in V_i$ for $i = 1, \ldots, d$, $(x_1, \ldots, x_d)$ and $(y_1, \ldots, y_d)$ are adjacent in $G_d^{\times}$ if and only if $(x_i, y_i) \in E_i$ for $i = 1, \ldots, d$. In our case, the matrices F_r, $r = 0, \ldots, n-1$, are the adjacency matrices of the (so-called distance) graphs G_r with common vertex sets $V_r = V = \{0, \ldots, n-1\}$, and with edge sets defined by: for $i, j \in \{0, \ldots, n-1\}$, $(i, j) \in E_r$ only when $|i - j| = r$. Then, $V_d^{\times} = \Gamma$, and for each $\alpha \in \Gamma$ we may define a product $G_d^{\times}(\alpha)$ of the graphs $G_{\alpha(i)}$ as above. It is known that $G_d^{\times}(\alpha)$ has adjacency matrix $F_\alpha^{\otimes}$ (Cvetković et al. [7, Theorem 2.21]). Thus, for any subset U of Γ, the union of the graphs $G_d^{\times}(\alpha)$ has adjacency matrix $A_U = \sum_{\alpha \in U} F_\alpha^{\otimes}$. Proposition 1 gives the case $U = \Gamma_h$.

Call two sequences (β, δ) *h-neighbors* if the sequence α defined in (10) is in Γ_h. This definition of neighbors—based on the Euclidean distance between points—is natural in some contexts, but in others a neighborhood structure based, say, on the L_1-norm (the length of the shortest walk connecting β to δ) may be more appropriate. The observation in the previous paragraph makes it straightforward to extend the results to follow to this case (and to neighborhood structures defined by other L_p-norms), but we omit the details.

2.2. Diagonal part

The matrices D_h in (3) are diagonal matrices with diagonal elements $D_h(\alpha)$ equal to the number of *h-neighbors* of α. In dimension $d = 1$ define, for each $r = 0, \ldots, n - 1$, the diagonal matrix M_r with ith diagonal element the ith row sum of F_r, and then define, for $\alpha \in \Gamma$,

$$M_\alpha^{\otimes} = M_{\alpha(1)} \otimes M_{\alpha(2)} \otimes \cdots \otimes M_{\alpha(d)}. \tag{11}$$

It is straightforward to prove:

Proposition 2. $D_h = \sum_{\alpha \in \Gamma_h} M_\alpha^{\otimes}$.

Notice that $tr[D_h]$ is the total number of non-zero elements in A_h, so that $tr[D_h] = 2N_h$. We have now established:

Theorem 1. *The spatial design matrix at distance $\sqrt{h}$ is given by*

$$L_h = \sum_{\alpha \in \Gamma_h} (M_\alpha^\otimes - F_\alpha^\otimes), \tag{12}$$

where $M_\alpha^\otimes$ and $F_\alpha^\otimes$ are as defined in (11) *and* (8).

The above expressions for the matrices A_h, D_h, and L_h involve summing over the set Γ_h. We next examine this set more closely, and give formulae for these matrices that do not involve Γ_h.

2.3. Generating functions

Since h must be a sum of squares of d of the integers $(0, 1, \ldots, n-1)$, not all values of $h \leqslant d(n-1)^2$ are feasible. This is so even when $d \geqslant 4$, notwithstanding Lagrange's four-square theorem [11, Section 20.5], because no term in the decomposition of h can exceed $(n-1)^2$. Thus, Γ_h in Proposition 1 can be empty, and in that case we define A_h, D_h and L_h to be zero matrices.

The values of h that yield non-vanishing matrices L_h can be read off from the expansion of the polynomial

$$(1 + t + t^4 + \cdots + t^{r^2} + \cdots + t^{(n-1)^2})^d = \sum_{h=0}^{d(n-1)^2} m_h t^h, \tag{13}$$

in which the coefficient m_h is evidently the number of ways in which h can be expressed as a sum of squares of d of the integers $(0, 1, \ldots, n-1)$, i.e., $m_h = |\Gamma_h|$ is the number of h-neighbors of the origin. Except for the restriction $h \leqslant d(n-1)^2$, the m_h evidently depend on d but not directly on n. Letting $f_n(t) = \sum_{r=0}^{n-1} t^{r^2}$, and using Wilf's [20] notation, we may write

$$m_h = [t^h](f_n(t))^d, \tag{14}$$

where $[t^h]$ means "the coefficient of t^h in the expansion of the following function in powers of t". Note that $[t^h]$ is identical to the operator $(h!)^{-1}(\partial/\partial t)^h|_{t=0}$, and, as an operator, is therefore linear. A cumbersome formula for the m_h can be deduced from (14), but using a modern symbolic computing package it is a simple matter to compute m_h from (14) without having to rely on such formulae.

Similarly, letting $b_n(t) = \Sigma_{r=0}^{n-1} t^{r^2} x_r$, where the x_i are labels for the integers $0, 1, \ldots, n-1$, obeying the usual rules of multiplication, we see that, from the formal expansion of $(b_n(t))^d$,

$$[t^h](b_n(t))^d = \sum_{\alpha \in \Gamma_h} \left\{ \prod_{i=1}^{d} x_{\alpha(i)} \right\}. \tag{15}$$

Thus, the sequence α belongs to Γ_h only if the product $\prod_{i=1}^{d} x_{\alpha(i)}$ appears on the right in (15).

The key to obtaining a simple representation for the matrices A_h, D_h, and hence L_h, is to notice that the scalar generating function $(b_n(t))^d$ can be generalized in such a way that, when expanded, the coefficient of t^h is precisely A_h. To see this, define the matrix

$$B_n(t) = \sum_{r=0}^{n-1} t^{r^2} F_r, \tag{16}$$

an $n \times n$ Toeplitz matrix with (i, j) element $t^{(i-j)^2}$. By direct expansion of the dth Kronecker power $B_n^{\otimes}(t) = \bigotimes_1^d B_n(t)$, it is clear that A_h is the coefficient of t^h in the expansion of $B_n^{\otimes}(t)$ in powers of t. That is,

$$A_h = [t^h] B_n^{\otimes}(t). \tag{17}$$

Similarly, letting

$$C_n(t) = \sum_{r=0}^{n-1} t^{r^2} M_r \tag{18}$$

and $C_n^{\otimes}(t) = \bigotimes_1^d C_n(t)$, we see that

$$D_h = [t^h] C_n^{\otimes}(t). \tag{19}$$

We therefore have the simple generating-function representation for L_h given in:

Theorem 2. *The spatial design matrix at distance $\sqrt{h}$ is given by*

$$L_h = [t^h](C_n^{\otimes}(t) - B_n^{\otimes}(t)). \tag{20}$$

These results evidently do not require knowledge of Γ_h: it is built in to the generating function. On the other hand, the matrices appearing in these representations of A_h, D_h and L_h are $N \times N$, and likely to be high-dimensional, so it might seem that these results would be of little practical value. On the contrary, we will see in the next section that they provide both analytically and computationally convenient information about the statistics q_h and q_h^* discussed in the Introduction, and hence about the properties of the variogram estimator $2\hat{\gamma}(h)$. Before doing so we note some further implications of these results.

It is clear that, if $\alpha \in \Gamma_h$, so is every permutation of the elements of α. Thus, Γ_h must be a union of one or more orbits in Γ under the action of the symmetric group S_d (the group of permutations of d objects). A set of orbit representatives is provided by the set $\Omega = \Omega(d, n)$ of *non-decreasing* sequences $\omega = (\omega(1), \ldots, \omega(d)) \in \Gamma$, with $\omega(1) \leqslant \omega(2) \leqslant \cdots \leqslant \omega(d)$. Let $\Omega_h = \{\omega \in \Omega : \|\omega\|^2 = h\}$, and, for $j = 0, \ldots, n-1$, $\omega \in \Omega$, let $k_\omega(j)$ denote the multiplicity of j in ω, so that $\sum_{j=0}^{n-1} k_\omega(j) = d$, and write $v(\omega) = \prod_{j=0}^{n-1} k_\omega(j)!$, with, as usual, $0! \equiv 1$.

With this notation it is easy to see that $m_h = d! \sum_{\omega \in \Omega_h} (v(\omega))^{-1}$, and since $\Gamma_h = \{\sigma\omega : \omega \in \Omega_h, \sigma \in S_d\}$, where $\sigma\omega$ denotes the permutation σ of ω, we have that

$$A_h = \sum_{\omega \in \Omega_h} \frac{1}{v(\omega)} F^*_\omega,$$

where $F^*_\omega = \sum_{\sigma \in S_d} F^{\otimes}_{\sigma\omega}$ is a symmetric function of the matrices $F_{\omega(1)}, \ldots, F_{\omega(d)}$. By an obvious extension of this argument to the off-diagonal part, and setting $M^*_\omega = \sum_{\sigma \in S_d} M^{\otimes}_{\sigma\omega}$, we can state:

Theorem 3. *The spatial design matrix at distance $\sqrt{h}$ is given by*

$$L_h = \sum_{\omega \in \Omega_h} \frac{1}{v(\omega)} (M^*_\omega - F^*_\omega). \tag{21}$$

For many values of h Eq. (15) reveals that Γ_h consists of a single orbit, which is to say that Ω_h has a single element, say ω_h. In that case $m_h = d!/v(\omega_h)$, and Theorem 3 gives the very simple result that $L_h = (v(\omega_h))^{-1}(M^*_{\omega_h} - F^*_{\omega_h})$. In the example following Proposition 1, for instance, $h = 1$, $\omega_1 = (0, .., 0, 1)$ and $v(\omega_1) = (d-1)!$.

Using these results we may also obtain the following generalization and simplification of Lemma 6.1 and Theorem 6.1 in GGS, which give upper bounds on the largest eigenvalues of L_h and A_h (for sets Ω_h with low cardinality), and hence upper bounds for the normalized statistics $z'L_hz/z'z$ and $z'A_hz/z'z$.

Lemma 1. *Let λ_h and μ_h denote the largest eigenvalues of A_h and L_h, respectively, and let $u_h = d! \sum_{\omega \in \Omega_h} \frac{2^{d-k_\omega(0)}}{v(\omega)}$. Then $\lambda_h \leqslant u_h$ and $\mu_h \leqslant 2u_h$.*

Proof. Let $g_h = \max_{\alpha \in \Gamma} D_h(\alpha)$ denote the maximum number of h-neighbors for any point in the grid Γ. The number m_h is the number of h-neighbors of the origin, so that $g_h \geqslant m_h$. Under the condition that no sequence $\alpha \in \Gamma_h$ contains an element $\alpha(i) > n/2$, we have $g_h = u_h$. To see this, suppose first that Ω_h contains just the single sequence ω_h. If $k_{\omega_h}(0) = 0$, $g_h = 2^d m_h$ because, under the stated condition, $\max_{\alpha \in \Gamma} D_h(\alpha)$ occurs at a sequence α for which the h-neighbors in all 2^d directions enter $D_h(\alpha)$, and m_h counts just the h-neighbors β in the direction for which the vector $\beta - \alpha$ has only positive components. If $k_{\omega_h}(0) > 0$, only $2^{d-k_{\omega_h}(0)}$ distinct directions are needed. Repeating the argument for each $\omega \in \Omega_h$ proves the claim $g_h = u_h$. Finally, when the condition that no $\alpha(i)$ exceeds $n/2$ is dropped, it is clear that $g_h \leqslant u_h$. The assertions $\lambda_h \leqslant u_h$, $\mu_h \leqslant 2u_h$ follow by Gershgorin's theorem (see [16]). □

If Ω_h contains only the single sequence ω_h, which contains only one non-zero term (so Γ_h contains only what GGS call "non-diagonal directions"), the matrices in the sum $\sum_{\alpha \in \Gamma_h} F^{\otimes}_\alpha$ *are* pairwise commutative, so the eigenvalues of A_h are simple functions of those of the single matrix F_r ($r = \sqrt{h}$) involved. Under the same condition, $L_h = ((d-1)!)^{-1} L^*_{\omega_h}$, with $L^*_{\omega_h} = \sum_{\sigma \in S_d} L^{\otimes}_{\sigma\omega_h}$, which is also a sum of pairwise commutative matrices. Thus, as

GGS note in Lemma 5.1, in the case of non-diagonal directions the eigenvalues of L_h are simple functions of those of the matrix $(M_{\sqrt{h}} - F_{\sqrt{h}})$.

The necessary and sufficient conditions required to ensure pairwise commutativity of the summands in Theorem 3 are that Ω_h contains only the single sequence ω_h, and ω_h contains no more than one (possibly repeated) non-zero integer. Note that ω_h may correspond to what GGS would call "diagonal directions", and that these conditions are always satisfied for $h = 1, 2, 3$ (for any $d \geqslant h$), but otherwise clearly hold only for special values of h.

3. Applications

In this section, we use the results established above to study the properties of the statistics $q_h^* = z'A_h z$ and $q_h = z'L_h z$. We consider first the case in which $z \sim N(0, I_N)$, but in Section 3.2 show how our earlier results can be used to deal with the more general case $z \sim N(0, \Sigma)$, assuming the process is second-order stationary and isotropic.

3.1. *Properties of the quadratic forms q_h^* and q_h*

Under the assumption $z \sim N(0, I)$, the distributions of the quadratic form $q = z'Az$, and its normalized form $\bar{q} = z'Az/z'z$, can certainly be obtained (see [14] for the former, and [13] for the latter), but both are sufficiently complicated as to inhibit their use for practical study of, and/or tabulation of, the distribution. On the other hand, it is well known that the cumulants of $q = z'Az$ under the assumption $z \sim N(0, \Sigma)$ are given by

$$\kappa_p = 2^{p-1}(p-1)!tr[(A\Sigma)^p], \;\; p = 1, 2, \ldots \tag{22}$$

(see [15] Chapter 3 for the definition of cumulants, and Chapter 15 for the result given in Eq. (22)). The results in Section 2 allow these cumulants to be computed quite straightforwardly when $\Sigma = I_N$ and the matrix A in (22) is either A_h or L_h. These results are given next. First, for comparison, we summarize the properties of the analogue of q_h^* for the case $d = 1$.

In the case $d = 1$ the properties of the statistics $Q_r^* = y'F_r y$, $r = 1, \ldots, n-1$, with $y \sim N(0, I_n)$, have been extensively studied. The following Lemma summarizes some elementary properties of the statistics Q_r^*, all of which are either given in, or are easily deduced from, the comprehensive results in [2]:

Lemma 2. *For $r = 1, \ldots, n-1$, let $Q_r^* = y'F_r y$, and assume that $y \sim N(0, I_n)$. Then:*

$$E(Q_r^*) = tr[F_r] = 0, \;\; and$$

$$var(Q_r^*) = 2\,tr[F_r^2] = 2\,tr[M_r] = 4(n-r).$$

All odd cumulants of Q_r^* vanish, so the density of Q_r^* is symmetric about zero, and for $r_1 \neq r_2$, $Q_{r_1}^*$ and *$Q_{r_2}^*$ are uncorrelated.*

3.1.1. *Properties of the q_h^**

With $A = A_h$ and $\Sigma = I_N$ in (22) we obtain the cumulants, $\kappa_{p,h}^*$, of q_h^*. Much of Lemma 2 generalizes easily to this case:

Lemma 3. *For $h \geqslant 1$, any $d \geqslant 1$, and $z \sim N(0, I_N)$,*

$$E(q_h^*) = tr[A_h] = 0,$$

$$var(q_h^*) = 2\,tr[A_h^2] = 2\,tr[D_h],$$

and, for $h_1, h_2 \geqslant 1,\ h_1 \neq h_2$, $q_{h_1}^*$ and $q_{h_2}^*$ are uncorrelated.

Proof. The first two cumulants are straightforward. To show that $cov(q_{h_1}^*, q_{h_2}^*) = 2\,tr[A_{h_1}A_{h_2}] = 0$, consider a diagonal element of $A_{h_1}A_{h_2}$:

$$(A_{h_1}A_{h_2})_{\alpha,\alpha} = \sum_{\beta \in \Gamma} (A_{h_1})_{\alpha,\beta}(A_{h_2})_{\beta,\alpha}\ \alpha \in \Gamma.$$

The product $(A_{h_1})_{\alpha,\beta}(A_{h_2})_{\beta,\alpha}$ vanishes unless both $\|\alpha - \beta\|^2 = h_1$ and $\|\beta - \alpha\|^2 = h_2$, which is impossible. Hence, for each $\alpha \in \Gamma$, every term in the sum on the right here vanishes. $\square$

Now, with the help of the generating function $C_n^{\otimes}(t)$ for D_h, it is straightforward to obtain a generating function for the variances $var(q_h^*)$, since

$$\begin{aligned} var(q_h^*) &= 2\,tr[D_h] = 2\,tr\{[t^h]C_n^{\otimes}(t)\} \text{ (using (19))} \\ &= 2[t^h]tr\{C_n^{\otimes}(t)\} \\ &= 2[t^h](tr(C_n(t))^d. \end{aligned} \tag{23}$$

The last step here follows from a standard property of the trace operator for Kronecker products, and the penultimate step from the fact that the operator $[t^h]$ commutes with the trace operator. Noting that $tr[M_0] = n$, and $tr[M_r] = 2(n-r),\ r = 1, \ldots, n-1$, it follows from the definition of $C_n(t)$ that

$$tr(C_n(t)) = (n + 2(n-1)t + \cdots + 2(n-r)t^{r^2} + \cdots + 2t^{(n-1)^2}). \tag{24}$$

Since $2N_h = tr[D_h]$, these formulae provide simple and efficient methods for computing the values N_h : setting $g_n(t) = tr(C_n(t))$ we have

$$2N_h = [t^h](g_n(t))^d. \tag{25}$$

In general, for $d > 1$, the density of q_h^* is not symmetric about zero. The analogue of the symmetry result for the case $d = 1$ in Lemma 2 is the weaker result given in:

Lemma 4. *If ph is odd $tr[A_h^p] = 0$ (independently of d). Hence, for h odd, the distribution of q_h^* (and also its normalized form $\bar{q}_h^* = q_h^*/z'z$) is symmetric about zero.*

Proof. Consider a diagonal element of A_h^p:

$$(A_h^p)_{\alpha,\alpha} = \sum_{\beta_1,\beta_2,\ldots,\beta_{p-1} \in \Gamma} (A_h)_{\alpha,\beta_1}(A_h)_{\beta_1,\beta_2} \cdots (A_h)_{\beta_{p-2},\beta_{p-1}}(A_h)_{\beta_{p-1},\alpha},\ \alpha \in \Gamma.$$

This is non-zero only if

$$\|\alpha - \beta_1\|^2 = \|\beta_1 - \beta_2\|^2 = \cdots = \|\beta_{p-1} - \alpha\|^2 = h.$$

Expanding each term $\|\beta_i - \beta_{i+1}\|^2$ as $\|\beta_i\|^2 + \|\beta_{i+1}\|^2 - 2\langle\beta_i, \beta_{i+1}\rangle$ and adding the p terms gives (with $\beta_0 = \beta_p = \alpha$):

$$2\left(\|\alpha\|^2 + \sum_{i=1}^{p-1}\|\beta_i\|^2 - \sum_{i=0}^{p-1}\langle\beta_i, \beta_{i+1}\rangle\right) = ph.$$

The left side is certainly an even integer, so when ph is odd we obtain a contradiction. Thus, when ph is odd, every term in the expression above for $(A_h^p)_{\alpha,\alpha}$ vanishes, for all $\alpha \in \Gamma$, implying $tr[A_h^p] = 0$. $\square$

The following result is also of some interest:

Lemma 5. *For $d = 2$ and every $h \geqslant 1$, $tr[A_h^3] = 0$.*

Proof. The diagonal element of A_h^3 labelled by (α, α) is given by

$$(A_h^3)_{\alpha,\alpha} = \sum_{\beta,\gamma\in\Gamma}(A_h)_{\alpha,\beta}(A_h)_{\beta,\delta}(A_h)_{\delta,\alpha}$$

and is non-zero only if there are $\beta, \delta \in \Gamma$ satisfying

$$\|\alpha - \beta\|^2 = \|\beta - \delta\|^2 = \|\delta - \alpha\|^2 = h.$$

This equation asserts that (α, β, δ) must be the vertices of an equilateral triangle in R^2, and it is well-known that there is no equilateral triangle with vertices in a two-dimensional integer grid (see, for instance, [3]), so this condition cannot be met for any α if $d = 2$. $\square$

Hence, if $d = 2$, $\kappa_{3,h}^* = 8tr[A_h^3] = 0$. The analogous result for dimensions $d > 2$ fails because in that case there are equilateral triangles in a uniform grid.

3.1.2. Properties of the q_h

We now deal with the case $A = L_h$ and $\Sigma = I_N$ in (22). Since $L_h l_N = 0$ (where l_N is an $N \times 1$ vector of ones), the results to follow continue to hold under the assumption that $z \sim N(\mu l_N, I_N)$, i.e., that the $Z(\alpha)$ have an unknown constant mean μ. We have, in either case, for the cumulants of q_h, $\kappa_{p,h} = 2^{p-1}\Gamma(p)tr[L_h^p]$, $p = 1, 2, \ldots$. Thus:

Lemma 6. *When $z \sim N(\mu l_N, I_N)$,*

$$E(q_h) = tr[L_h] = tr[D_h] = 2N_h \tag{26}$$

and

$$var(q_h) = 2\,tr[L_h^2] = 2(tr[D_h^2] + tr[D_h]). \tag{27}$$

The result for the variance uses the facts that $tr[D_h A_h] = 0$ and $tr[A_h^2] = tr[D_h]$. The computation of $tr[D_h]$ has been discussed above, and we can compute the term $tr[D_h^2]$ from the formula:

$$tr[D_h^2] = tr[[t^h][s^h]C_n^{\otimes}(t)C_n^{\otimes}(s)] = [(ts)^h](tr[C_n(t)C_n(s)])^d.$$

Thus:

$$var(q_h) = 2\{[(ts)^h](tr[C_n(t)C_n(s)])^d + 2N_h\}. \tag{28}$$

From the definition of $C_n(t)$, $tr[C_n(t)C_n(s)] = \sum_{r_1,r_2=0}^{n-1} t^{r_1^2} s^{r_2^2} tr[M_{r_1} M_{r_2}]$, and it is easy to check that $tr[M_0^2] = n$ and, for $1 \leqslant r_1 \leqslant r_2 \leqslant n-1$,

$$tr[M_{r_1} M_{r_2}] = \begin{cases} 4(n-r_2) - 2r_1 & \text{if } r_1 + r_2 \leqslant n, \\ 2(n-r_2) & \text{otherwise.} \end{cases} \tag{29}$$

Thus, we again have a simple generating function for the variances of the statistics q_h, and hence for the variance of the variogram estimator in the "null" case ($\Sigma = I_N$) (see Section 3.3 below).

Higher-order cumulants and product cumulants (e.g., covariances) for both the q_h^* and the q_h can be obtained by obvious extensions of these methods. For instance,

$$tr[A_h^p] = [(t_1 \cdots t_p)^h](tr[B_n(t_1) \cdots B_n(t_p)])^d \tag{30}$$

and

$$cov(q_{h_1}, q_{h_2}) = 2\,tr[L_{h_1} L_{h_2}] = 2\,tr[D_{h_1} D_{h_2}] = 2[t^{h_1}][s^{h_2}](tr[C_n(t)C_n(s)])^d. \tag{31}$$

The generating functions in these expressions may, of course, simplify (as above), and this reduces the computational problem considerably. We leave other such extensions to the reader.

3.2. *Second-order stationary isotropic processes*

Under the assumption that the process is second-order stationary and isotropic—which is stronger than the intrinsic stationarity assumption mentioned in the introduction (see [6])—we have, as an obvious consequence of Eq. (17):

Proposition 3. *If the process $\{Z(\alpha); \alpha \in \Gamma\}$ is second-order stationary and isotropic, its covariance matrix Σ has the representation*

$$\Sigma = \sum_{h \in H} c(h) A_h, \tag{32}$$

where H is a some set of values of h containing zero (recall that $A_0 = I_N$), and the coefficients $\{c(h); h \in H\}$ must be such that Σ is positive definite. Thus, from (17), $\Sigma = [S_H(t)]B_n^{\otimes}(t)$, where

$$[S_H(t)] = \sum_{h \in H} c(h)[t^h]. \tag{33}$$

The operator $[S_H(t)]$ constructs a linear combination, with parameters $c(h)$, of the coefficients of the powers t^h, $h \in H$, that occur in the expansion of the function to which it is applied. Like the $[t^h]$ themselves, $[S_H(t)]$ is clearly linear. If we now assume that $z \sim N(0, \Sigma)$, with Σ as in (32), and take $h > 0$, we easily see that:

$$E(q_h^*) = tr[A_h\Sigma] = \sum_{k\in H} c(k)tr[A_hA_k] = \begin{cases} c(h)tr[D_h] & \text{if } h \in H, \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

And (since $tr[A_h] = tr[D_kA_h] = 0$),

$$\begin{aligned} E(q_h) &= tr[L_h\Sigma] = \sigma^2 tr[D_h] - \sum_{k\in H\backslash\{0\}} c(k)tr[A_hA_k] \\ &= \begin{cases} \{\sigma^2 - c(h)\}tr[D_h] & \text{if } h \in H, \\ \sigma^2 tr[D_h] & \text{otherwise,} \end{cases} \end{aligned} \tag{35}$$

where we have put $c(0) = \sigma^2$. Since, under these assumptions, $\gamma(h) = \sigma^2 - c(h)$, this shows that $2\hat{\gamma}(h) = q_h/N_h$ is an unbiased estimator of the true variogram $2\gamma(h)$, for all $h > 0$, as is well-known [6]. Obviously, to compute the unbiased estimator $2\hat{\gamma}(h)$ one needs to know the correct scale factor N_h, and this has hitherto been unavailable for the isotropic case in general; Eq. (25) gives a simple general procedure for computing it, generalizing the special case given in Lemma 7.1 in GGS.

The variances and covariances of the statistics q_h^* and q_h for several values of h are often needed in applications. For instance, the entire covariance matrix of a vector of statistics q_h at a set of values of h is required for variogram fitting by generalized least squares [9,6, Section 2.6.2], and this has previously been unavailable for the isotropic case. The covariances cannot easily be written down in closed form, but when Σ has the form (32) are easily represented in generating function form using the operators $[S_H(t)]$ defined in (33). Thus we easily obtain:

Lemma 7. *Suppose $z \sim N(0, \Sigma)$, with Σ of the form (32). Then, for any $h_1 \geqslant h_2$:*

$$\begin{aligned} cov(q_{h_1}^*, q_{h_2}^*) &= 2\,tr[A_{h_1}\Sigma A_{h_2}\Sigma] \\ &= 2[s_1^{h_1}][s_2^{h_2}][S_H(t_1)][S_H(t_2)]v_n^d(s_1, s_2, t_1, t_2) \end{aligned} \tag{36}$$

and

$$\begin{aligned} cov(q_{h_1}, q_{h_2}) &= 2\,tr[L_{h_1}\Sigma L_{h_2}\Sigma] \\ &= 2[s_1^{h_1}][s_2^{h_2}][S_H(t_1)][S_H(t_2)]V_n^d(s_1, s_2, t_1, t_2), \end{aligned} \tag{37}$$

where

$$v_n^d(s_1, s_2, t_1, t_2) = (tr[B_n(s_1)B_n(t_1)B_n(s_2)B_n(t_2)])^d \tag{38}$$

and

$$\begin{aligned} V_n^d(s_1, s_2, t_1, t_2) &= v_n^d(s_1, s_2, t_1, t_2) + (tr[C_n(s_1)B_n(t_1)C_n(s_2)B_n(t_2)])^d \\ &\quad -2(tr[C_n(s_2)B_n(t_2)B_n(s_1)B_n(t_1)])^d. \end{aligned} \tag{39}$$

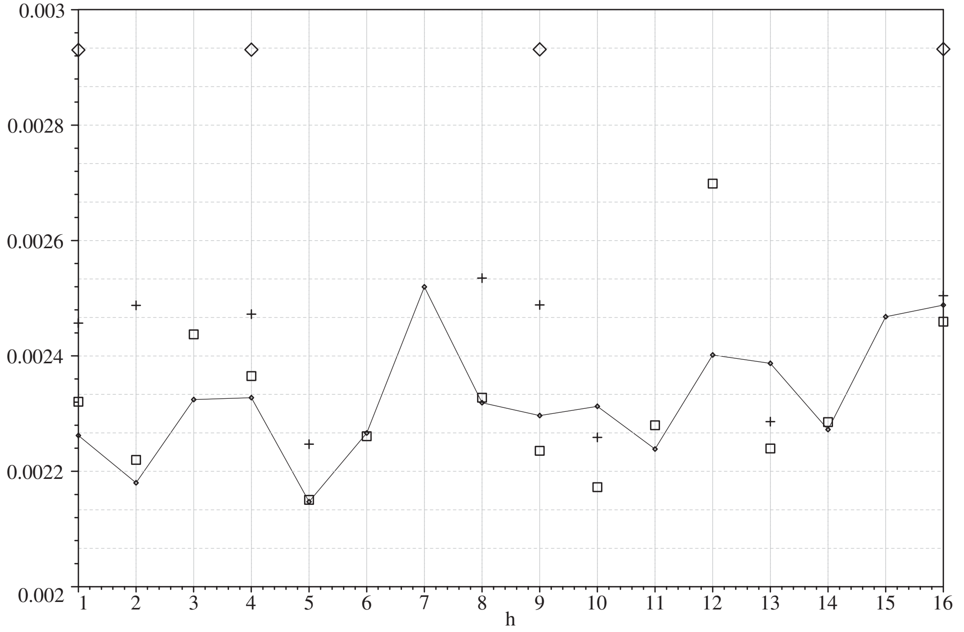

Fig. 1. The variance of the classical estimator $2\hat{\gamma}(h)$ as a function of h and d : $d = 1$ (diamond), 2 (cross), 3 (square), 4 (line); $N = 2^{12}$, $\Sigma = I_N$.

Note that $cov(q^*_{h_1}, q^*_{h_2}) = 0$ when $h_1 \neq h_2$ and $h_1, h_2 \notin H$, and that the elements of the matrix defining $v^d_n(s_1, s_2, t_1, t_2)$ are positive. Thus, if the $c(h)$ in (32) are positive and non-decreasing in $|H|$, an increase in $|H|$ must increase $cov(q^*_{h_1}, q^*_{h_2})$. Extensions to higher-order cumulants are obvious, but, as in the case $\Sigma = I_N$, will entail a larger computational burden. Finally, we note that the approach used here can also be extended to the case where the precision matrix Σ^{-1}, rather than Σ itself, is a linear combination of the A_h.

3.3. *Properties of the classical variogram estimator*

The above results for q_h provide the tools for studying the properties of the classical variogram estimator for a second-order stationary and isotropic process under virtually any specification for the $c(h)$. We do not intend to study the detailed properties of the variogram estimator here, but will show that the above results can be used to study the properties of $2\hat{\gamma}(h)$ under a variety of specifications for the variogram $2\gamma(h)$ (for the intrinsically stationary, but non-isotropic case, see [5]).

We first consider the variance of $2\hat{\gamma}(h) = q_h/N_h$ as a function of h and d, assuming $\Sigma = I_N$. In Fig. 1 we plot $var(2\hat{\gamma}(h)) = var(q_h)/N_h^2$, computed using Eqs. (25) and (28), for $d = 1, 2, 3, 4$, and $h = 1, \ldots, 16$, with N held fixed at $N = 2^{12}$, so that, for $d = 1, 2, 3, 4$ we have $n = 2^{12}, 2^6, 2^4, 2^3$ respectively.

Fig. 1 shows that: (a) for each fixed dimension $d > 1$, the variance is quite volatile as h varies; and (b) the variance is not monotonic in d for fixed h (see for instance the value

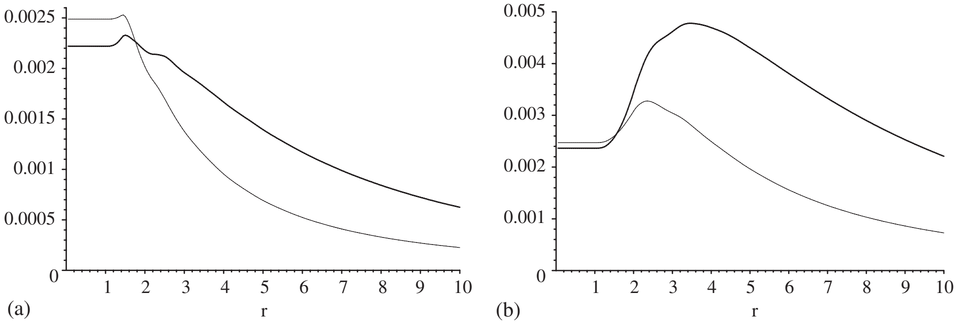

Fig. 2. The variance of the classical estimator $2\hat{\gamma}(h)$ when the variogram is spherical. In (a) $h = 2$, in (b) $h = 4$. The variance is plotted for many values of the range r from 0 to 10, $N = 2^{12}$, $d = 2$ (thin line), $d = 3$ (thick line).

$h = 9$). Thus, in contrast to Fig. 4 in GGS (where the variance could only be computed for "non-diagonal" directions), our results show that when "diagonal" directions are taken into account—as it is natural to do under the assumption of isotropy—$var(2\hat{\gamma}(h))$ is no longer monotonic either in d or in h. The volatility and non-monotonicity of the variances is attributable to variation in N_h, m_h, and the structure of Ω_h as h varies. The explanation is purely number theoretic: the number of decompositions of a particular h as a sum of squares is not related in any simple way to the values n and d.

The variance of the classical variogram estimator when Σ is of the form (32) can be computed using (37) with $h_1 = h_2$. Using this formula, one can study the behavior of $var(2\hat{\gamma}(h))$ under various specifications for the true variogram $2\gamma(h)$, i.e., of the $c(h)$ in (32). In Fig. 2 we plot the variances for the case of a spherical variogram with sill 1, nugget 0 and range r, so that the $c(h)$ in (32) are given by

$$c(h) = c(h, r) = \begin{cases} 1 - (3\sqrt{h}/r + (\sqrt{h}/r)^3)/2 & \text{if } 0 \leqslant h \leqslant r^2, \\ 0 & \text{if } h > r^2. \end{cases} \tag{40}$$

The value of N is kept fixed, as above, at $N = 2^{12}$. We plot the variances for $d = 2$ and $d = 3$ as a function of the range r (the variogram is not valid for $d > 3$). In Fig. 2(a) we display the results for $h = 2$ (note that this is a diagonal direction in the sense of GGS—for any d), and in Fig. 2(b) for $h = 4$. The corresponding figure for $h = 1$ is equivalent to Fig. 7 in GGS, which was produced by simulation for $N = 2^8$ (note that GGS appear to have omitted a factor 2).

In Fig. 3 we repeat this exercise for the case of an exponential variogram with sill 1, nugget 0 and (practical) range r, so that the $c(h)$ in (32) are given by

$$c(h) = c(h, r) = \exp\{-3\sqrt{h}/r\}, \ h \geqslant 0. \tag{41}$$

In this case, all feasible values of h will appear in Eq. (32), presenting a much larger computational task for the evaluation of $var(2\hat{\gamma}(h))$. Nevertheless, by exploiting the structure of the generating function (39) to streamline the computation, the variances can be computed efficiently. In Fig. 3(a) we plot the variances as a function of r for $h = 2$, and in Fig. 3(b)

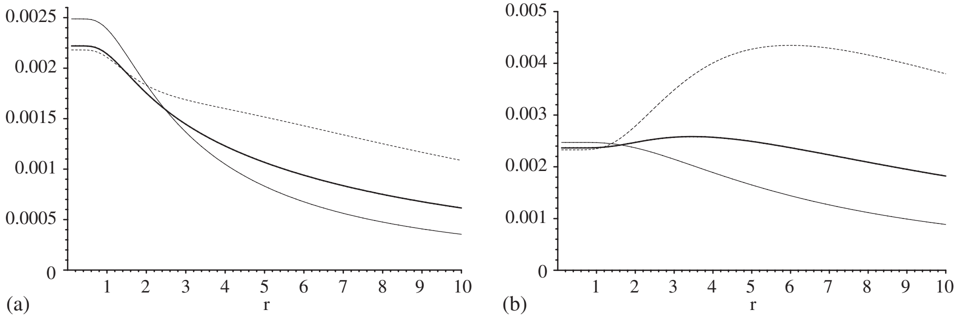

Fig. 3. The variance of the classical estimator $2\hat{\gamma}(h)$ when the variogram is exponential. In (a) $h = 2$, in (b) $h = 4$. The variance is plotted for many values of the (practical) range r from 0 to 10, $N = 2^{12}$, $d = 2$ (thin line), $d = 3$ (thick line), $d = 4$ (dashed line).

those for $h = 4$, in both cases for $d = 2$, 3, and 4 (the exponential variogram is valid for all d).

With a fixed number, N, of $i.i.d.$ observations, we expect the variance to decrease, at least for small h ($h \leqslant N^{\frac{1}{d}}$) as d increases, because the number of pairs of points available to estimate $2\gamma(h)$ (for fixed h) cannot decrease as d increases, and usually increases. But, as dependence in the data increases, or h increases, one anticipates that this effect might be overturned. Both Figs. 2 and 3 show that these expectations are correct: the variances are not monotonic in r, sometimes increasing with r initially, then decreasing. And the non-monotonicity is more pronounced for larger h, and for the case of a spherical variogram. Note that the lack of smoothness for low values of r evident in Fig. 2 arises because the spherical variogram itself is not smooth. For sufficiently large values of r—the values most likely to be used in applications—the variance for fixed h is increasing in d for both variograms—as suggested by GGS.

Of course, the usefulness of Lemma 7 is in providing a means to compute $var(2\hat{\gamma}(h))$ (and covariances) exactly in applications. For the exponential this is not a trivial computation, because as we note above, $c(h) \neq 0$ for all feasible values of h, so that $[S_H(t)]$ in (33) contains all feasible values. In practice, however, perfectly satisfactory accuracy can be achieved by truncating the $c(h, r)$ at some point.

4. Concluding remarks

We have provided simple formulae and generating functions for the spatial design matrices implicitly defined by quadratic forms that arise in the analysis of isotropic spatial models on uniform grids, extending and simplifying the results in [9,10]. Such models are a natural generalization of familiar time series models—the one-dimensional case—and the structural results we have derived reflect this relation. These results show that in general these matrices are sums of non-commuting matrices—Kronecker products of their counterparts for the one-dimensional case—and hence that their eigenvalues are unlikely to be expressible in terms of those of the summands.

Fortunately, to study the properties of the associated quadratic forms the eigenvalues themselves are not needed: the generating functions for the matrices themselves induce generating functions for their cumulants. We provide detailed results on the means, variances and covariances of these statistics. As an important application of these results, we give simple formulae for the normalizing constant needed to produce an unbiased estimator of the variogram, and, assuming second-order stationarity, the covariance matrix needed to implement generalized least squares procedure for variogram estimation (see [6, Chapter 6]). Finally, we briefly study some properties of the classical variogram estimator for the cases of some popular choices of the actual variogram.

For the purposes of hypothesis testing the normalized statistics $\bar{q}_h^* = z'A_h z/z'z$ and $\bar{q}_h = z'L_h z/z'z$ are of greater interest. But since exact distribution theory for such statistics is difficult, various techniques for approximating the distributions based on just the low-order cumulants have been developed (see, for instance, [1,8,12]). Although we do not implement them here, the results in Section 3 make such techniques quite straightforward. It is easily seen that, under the assumption that $z \sim N(0, \sigma^2 I_N)$—usually the null hypothesis—the ratios $\bar{q}_h^*$ and $\bar{q}_h$ are independent of their denominator, so that the moments of the ratios are ratios of the moments. Hence the cumulant results for q_h^* and q_h given in Section 3 can also be used to study or approximate the properties of $\bar{q}_h^*$ and $\bar{q}_h$ under this assumption.

It is, of course, both analytically and computationally convenient if the eigenvalues, or good approximations to them, of L_h and A_h are known. One possible device for developing approximations in the case $d = 1$ is to replace the F_r by their circular counterparts (see [2, Chapter 6.5]), and our results allow that approach to be adapted to higher dimensional cases straightforwardly. We will report our work on that subject elsewhere.

Acknowledgment

We thank two anonymous referees for helpful comments on an earlier version of the paper. FM acknowledges support from ESRC grant No. R42200134323.

References

[1] M.M. Ali, Durbin–Watson generalized Durbin–Watson tests for autocorrelations and randomness, J. Bus. Econom. Statist. 5 (1987) 195–203.

[2] T.W. Anderson, The Statistical Analysis of Time Series, Wiley, New York, 1971.

[3] M.J. Beeson, Triangles with vertices on lattice points, Amer. Math. Mon. 99 (1992) 243–252.

[4] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. Roy. Statist. Soc. Ser. B 36 (1974) 192–236.

[5] N. Cressie, Fitting varogram models by weighted least squares, Math. Geol. 17 (1985) 563–586.

[6] N. Cressie, Statistics for Spatial Data, Wiley, New York, 1993.

[7] D.M. Cvetković, M. Doob, H. Sachs, Spectra of Graphs, Academic Press, New York, 1980.

[8] J. Durbin, G.S. Watson, Testing for serial correlation in least squares regression II, Biometrika 38 (1951) 159–178.

[9] M.G. Genton, Variogram fitting by generalized least squares using an explicit formula for the covariance structure, Math. Geol. 30 (1998) 323–345.

[10] D.J. Gorsich, M.G. Genton, G. Strang, Eigenstructures of spatial design matrices, J. Multivariate Anal. 80 (2002) 138–165.

[11] G.H. Hardy, E.M. Wright, An Introduction to the Theory of Numbers, fifth ed., Oxford University Press, Oxford, 1979.
[12] R.C. Henshaw, Testing single-equation least squares regression models for autocorrelated disturbances, Econometrica 34 (1996) 646–660.
[13] G.H. Hillier, The density of a quadratic form in a vector uniformly distributed on the n-sphere, Econometric Theory 17 (2001) 1–28.
[14] A.T. James, Distributions of matrix variates and latent roots derived from normal samples, Ann. Math. Statist. 35 (1964) 475–501.
[15] M.G. Kendall, A. Stuart, The Advanced Theory of Statistics, vol. 1, Distribution Theory, Griffin and Co., London, 1969.
[16] M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Dover, New York, 1969.
[17] B. Mohar, Some applications of Laplace eigenvalues of graphs, in: G. Hahn, G. Sabidussi (Eds.), Graph Symmetry: Algebraic Methods and Applications, vol. 497 of NATO ASI Series C, Kluwer, Dordrecht, 1997, pp. 227–275.
[18] P.A.P. Moran, Notes on continuous stochastic phenomena, Biometrika 37 (1950) 17–23.
[19] J. von Neumann, R.H. Kent, H.R. Bellinson, B.I. Hart, The mean square successive differences, Ann. Math. Statist. 12 (1941) 153–162.
[20] H.S. Wilf, generatingfunctionology, second ed., Academic Press Inc., New York, 1994.