The Korean Communications
in Statistics Vol. 12 No. 3, 2005
pp. 653-657

A Note on Eigenstructure of a Spatial Design Matrix in R^1

Hyoung-Moon Kim[1] and Pablo Tarazaga[2]

Abstract

Eigenstructure of a spatial design matrix of Matheron's variogram estimator in R^1 is derived. It is shown that the spatial design matrix in R^1 with $n/2 \le h < n$ has a nice spectral decomposition. The mean, variance, and covariance of this estimator are obtained using the eigenvalues of a spatial design matrix. We also found that the lower bound and the upper bound of the normalized Matheron's variogram estimator.

Keywords : Eigenvalue, Eigenvector, Kriging, Matheron's estimator

1. Introduction

Variogram estimation is an important step of spatial statistics since it determines the kriging weights. To determine these weights, an optimal linear spatial predictor from the data is commonly used and such spatial prediction is called kriging (Cressie, 1993). Matheron's classical variogram estimator of an intrinsic stationary spatial process, $\{Y(\boldsymbol{x}): \boldsymbol{x} \in D \subset R^d, d \ge 1\}$, is as follows (Cressie, 1993):

$$2\hat{\gamma}(\boldsymbol{h}) = \frac{1}{N_{\boldsymbol{h}}} \sum_{N(\boldsymbol{h})} (Y(\boldsymbol{x}_i) - Y(\boldsymbol{x}_j))^2, \quad \boldsymbol{h} \in R^d, \; N(\boldsymbol{h}) = \{(\boldsymbol{x}_i, \boldsymbol{x}_j): \boldsymbol{x}_i - \boldsymbol{x}_j = \boldsymbol{h}\},$$

where $N_{\boldsymbol{h}}$ is the cardinality of $N(\boldsymbol{h})$. This estimator can be expressed as a quadratic form,

$$2\hat{\gamma}(\boldsymbol{h}) = \boldsymbol{y}^t \frac{1}{N_{\boldsymbol{h}}} A(\boldsymbol{h})\, \boldsymbol{y}, \quad \boldsymbol{y} = (Y(\boldsymbol{x}_1), \ldots, Y(\boldsymbol{x}_n))^t \tag{1}$$

where the spatial design matrix, $A_{\boldsymbol{h}} = A(\boldsymbol{h})/N_{\boldsymbol{h}}$, is given by Genton(1998) and Gorsich et al. (2002).

To understand the properties of Matheron's variogram estimator in R^1, we explore the eigenstructure of a spatial design matrix in R^1. The spatial design matrix of Matheron's classical variogram estimator, $A_h = A(h)/(n-h)$ of size $n \times n$, is given in the unidimensional

1) Assistant Professor, Department of Applied Statistics, Konkuk University, Seoul 143-701, Korea
E-mail : hmkim@konkuk.ac.kr

2) Associate Professor, Department of Mathematics, Texas A&M University-Corpus Christi, Corpus Christi, TX 78412, USA

case (Genton, 1998) when data are regularly spaced:

$$A_h = \frac{1}{n-h} A(h) = \frac{1}{n-h} \begin{pmatrix} I_{n-h} & -I_{n-h} \\ -I_{n-h} & I_{n-h} \end{pmatrix}. \tag{2}$$

It is built by superposing identity matrices I_{n-h}, of size $(n-h)\times(n-h)$. If $h < n/2$, superposition means that two elements located at the same place are added. Note that if data are irregularly spaced, tolerance regions around h are often used (Cressie, 1993). There are three possible matrices depending on h. For example, the spatial design matrices of $n=4$ after removing $1/(n-h)$ are:

$$A(1) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad A(2) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}, \quad A(3) = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

Gorsich et al. (2002) explored the eigenstructures of spatial design matrices only for the case of $h < n/2$, so we will extend the results on the case of $n/2 \le h < n$ in R^1 since the possible range of h is 0 to n. When $h=2/n$, the spatial design matrix is a special case of Toeplitz matrix. In Section 2, the eigenstructure of a spatial design matrix is given. In particular, it is shown that the spatial design matrix in R^1 with $n/2 \le h < n$ has a nice spectral decomposition. In Section 3, the mean, variance, and covariance of Matheron's variogram estimator are provided using the results of Section 2. We also found that the lower bound and the upper bound of the normalized Matheron's variogram estimator. Finally, a brief conclusion is offered in Section 4.

2. Eigenstructure

In this section, our attention will be on the eigenstructure of the $n\times n$ matrix A_h defined at (2), where $n/2 \le h < n$. Note that $rank(A_h) = n-h$, due to the fact that we have an identity left upper block of size $n-h$ and the fact that the last $n-h$ rows are the opposite of the first $n-h$ ones (which imply that they are linearly dependent).

If we denote $A_h(:,i)$ as the i^{th} column of A_h, then it is clear that $A_h(:,h+i) = -A_h(:,i)$ for $i=1,\dots,n-h$. This implies that two equations, which are going to be very useful, are

$$A_h(:,i) + A_h(:,h+i) = 0 \text{ and} \tag{3}$$

$$A_h(:,i) - A_h(:,h+i) = 2A_h(:,i), \; i=1,\dots,n-h. \tag{4}$$

Now we define three families of vectors. If we denote e_i as the canonical vectors (1 in the i^{th} component and the rest zero), then we can define vectors $v_i = e_i + e_{i+h}$ and $u_i = e_i - e_{i+h}$, $i=1,\dots,n-h$. Finally we have the following result.

Lemma 1 Given an integer h such that $n/2 \le h < n$,

1) The vectors v_i, $i=1,\dots,n-h$, and the vectors e_i, $i=n-h+1,\dots,h$, are in the null space of A_h.

2) The vectors $u_i,\ i=1,\ldots,n-h,$ satisfy the following equation

$$A_h u_i = \frac{2}{n-h} u_i$$

Proof. First of all, observe that $A_h v_i$ is equivalent to $A_h(:,i)+A_h(:,h+i)$ given the structure of vector v_i. The sum is 0 because of equation (3). Using the same argument, we find that $A_h e_i = A_h(:,i),\ i=n-h+1,\ldots,h.$ Note that these columns are zero columns, so the result follows. Finally $A_h u_i = A_h(:,i)-A_h(:,h+i)$ and, using equation (4), we obtain $A_h u_i = \frac{2}{n-h} u_i$.

Therefore, eigenvalues of A_h are $2/(n-h)$ with multiplicity $n-h$ and 0 with multiplicity h. Corresponding eigenvectors are $u_i,\ i=1,\ldots,n-h,$ $v_i,\ i=1,\ldots,n-h$ and $e_i,\ i=n-h+1,\ldots,h,$ respectively.

Now we can define the following matrices,

$$Q=\left[\frac{v_1}{\sqrt{2}}\cdots\frac{v_{n-h}}{\sqrt{2}}\cdots e_{n-h+1}\cdots e_h \frac{u_1}{\sqrt{2}}\cdots\frac{u_{n-h}}{\sqrt{2}}\right],\quad n/2\le h< n,$$

and a diagonal matrix Λ whose diagonal elements are 0 for the first h entries and $2/(n-h)$ for the rest.

Lemma 2 The matrix A_h can be factorized as follows:

$$A_h = Q\Lambda Q^t.$$

Proof. This result is clear from the facts that 0 is the eigenvalues for the eigenvectors $v_i,\ i=1,\ldots,n-h$ and $e_i,\ i=n-h+1,\ldots,h.$ It is also straightforward to verify that the matrix Q is an orthonormal matrix.

So the matrix A_h has a very nice spectral decomposition.

3. Matheron's Variogram Estimator

The properties of Matheron's variogram estimator can be explained as functions of a spatial design matrix, A_h, and the actual covariance matrix of the data, . The mean and variance of this estimator can be obtained using all the eigenvalues of A_h and are described in the next Lemma. A Gaussian assumption can be generalized to elliptical distributions with kurtosis parameter 0(Genton, 2000), skew Gaussian distributions (Genton et al., 2001) or skew t distributions (Kim and Mallick, 2003).

Theorem 1 If $y\sim N_n(\mu,\Omega)$, where $\mu=\mu_y 1_n$, $\mu_y=E(y_i),\ i=1,\ldots,n$ and $\Omega=\sigma^2 I_n$, then the sample variogram estimator (1) with $A_h=A(h)/(n-h)$ and $n/2\le h< n,$ satisfies:

1) $E(y^t A_h y) = tr(A_h \Omega) = 2\sigma^2,$

2) $Var(y^t A_h y) = 2tr((A_h \Omega)^2) = 8\sigma^4/(n-h),$

and for $h_1 < h_2$

3) $Cov(y^t A_{h_1} y,\ y^t A_{h_2} y) = 2tr(A_{h_1} \Omega A_{h_2} \Omega) = 4\sigma^4/(n-h_1).$

Proof. In general, if $y \sim N_n(\mu, \Omega)$ then $E(y^t A_h y) = tr(A_h \Omega) + \mu^t A_h \mu$, $Var(y^t A_h y) = 2tr((A_h \Omega)^2) + 4\mu^t A_h \Omega A_h \mu$, and $Cov(y^t A_{h_1} y,\ y^t A_{h_2} y) = 2tr(A_{h_1} \Omega A_{h_2} \Omega) + 4\mu^t A_{h_1} \Omega A_{h_2} \mu$. Since $A_h 1_n = 0$ by the property of A_h, first equalities at 1), 2), and 3) are clear. This fact was commented by one of the referees. The sum of all eigenvalues of a matrix is the trace of the matrix and an eigenvalue of A^k is λ^k, where λ is an eigenvalue of A and $k \geq 1$ is an integer. The form for the covariance follows the same reasoning, but the trace of $A_{h_1} A_{h_2}$ is needed. This trace has been found in Genton(1998) as $2/(n-h_1)$, where $h_1 < h_2$ and $h_1 + h_2 \geq n$. Now the entire Theorem is proved.

Above results can be used at Generalized least squares method (Genton, 1998). One of the referees noted this potential application of Theorem. We note that the variance of the sample variogram estimator is an increasing function of lag h. So the variance of the sample variogram will go to the infinity as lag h goes to n. Furthermore we can easily find the lower bound and the upper bound of the normalized form of Matheron's estimator $y^t A_h y / y^t y$ using the eigenvalues of A_h. The normalization constant is in fact $c(0)$, the covariogram at lag 0. Suppose that the spatial process is second order stationary, so that $\gamma(h) = c(0) - c(h)$. The minimum (maximum) eigenvalue is the lower (upper) bound of the normalized quadratic form, respectively. That is $0 \leq y^t A_h y / y^t y \leq 2/(n-h)$. Moreover, there are well established bounds on normalized covariograms(Yaglom, 1987a; Yaglom, 1987b).

4. Conclusion

Eigenstructure of a spatial design matrix of Matheron's variogram estimator in R^1 with $n/2 \leq h < n$ was derived at Section 2. In particular, it was shown that the spatial design matrix in R^1 with $n/2 \leq h < n$ has a nice spectral decomposition. Using the eigenstructure of a spatial design matrix, we derived the mean, variance, and covariance of Matheron's variogram estimator and we found that the lower bound and the upper bound of the normalized quadratic form at Section 3.

Acknowledgments

The authors thank the editor and two referees for their constructive suggestions that led to significant improvements in the presentation of the article. The first author's work was supported by the faculty research fund of Konkuk University in 2004.

References

[1] Cressie, N., (1993). *Statistics for Spatial Data(rev. ed.),* Wiley, New York.

[2] Genton, M.G., (1998). Variogram fitting by generalized least squares using an explicit formula for the covariance structure. *Mathematical Geology.* Vol. 30, 323-345.

[3] Genton, M.G., (2000). The correlation structure of Matheron's classical variogram estimator under elliptically contoured distributions. *Mathematical Geology*, Vol. 32, 127-137.

[4] Genton, M.G., He, L. and Liu, X., (2001). Moments of skew-normal random vectors and their quadratic forms. *Statistics & Probability Letters*, Vol. 51, 319-325.

[5] Gorsich, D.J., Genton, M.G., and Strang, G., (2002). Eigenstructures of Spatial Design Matrices. *Journal of Multivariate Analysis*, Vol. 80, 138-165.

[6] Kim, H.-M., Mallick, B.K., (2003). Moments of random vectors with skew t distribution and their quadratic forms. *Statistics & Probability Letters*, Vol. 63, 417-423.

[7] Yaglom, A.M., (1987a). *Correlations Theory of Stationary and Related Random Functions. I. Basic Results*, Springer Series in Statistics, Springer-Verlag, Berlin/New York.

[8] Yaglom, A.M., (1987b). *Correlations Theory of Stationary and Related Random Functions. II. Supplementary Notes and References*, Springer Series in Statistics, Springer-Verlag, Berlin/New York.

[Received December 2004, Accepted October 2005]