A note on random fields forming conditional bases

Naveen Bansal\textsuperscript{a}, G.G. Hamedani\textsuperscript{a}, Hao Zhang\textsuperscript{b, *}

\textsuperscript{a}Department of Mathematics, Statistics, and Computer Science, Marquette University, P.O. Box 1881, Milwaukee, WI 53201-1881, USA
\textsuperscript{b}The Program in Statistics, Washington State University, P.O. Box 643144, Pullman, WA 99164, USA

Received February 1999; received in revised form March 2000

Abstract

We study properties of random fields that form conditional bases and their applications in spatial statistics. \textcopyright{} 2000 Elsevier Science B.V. All rights reserved

Keywords: Conditional basis; Random fields; Spatial statistics; Variogram

1. Introduction

Let \( \{ X_t: t \in D \} \) be a set of real-valued random variables on \( D \), a countable subset of \( \mathbb{R}^d \). For \( d = 2 \) or 3, \( \{ X_t: t \in D \} \) represents the spatial or spatio-temporal observations that are recorded at different locations or time, e.g., the ground water contamination near a nuclear fertility, the amount of rainfall at monitoring sites and time. An important problem is to predict \( X_s \), for an \( s \notin D \) based on \( \{ X_t: t \in D \} \). The linear prediction by the least-squares method is widely used due to its simple form.

As in many studies of random fields, we suppose \( X_t \in L^2(dP) \), \( \forall t \in D \) have zero means. Indeed, the means or trend surface can be properly removed by some standard techniques. When \( D \) is a finite set, the best linear predictor of \( X_s, s \notin D \), based on \( \{ X_t: t \in D \} \) can be written as a sum:

\[
\hat{X}_s = \sum_{t \in D} a(t,s)X_t.
\]  

(1)

When \( D \) is countable but contains infinitely many points, \( \hat{X}_s \) may not be written as a convergent series and there is even a need to define what the convergence is. Even in the case \( d = 1 \), the problem of expressing \( \hat{X}_s \) as a strongly convergent series is very challenging. Research has been carried out to express \( \hat{X}_s \) as a series that converges in a weaker sense. Rozanov (1963, Chapter II) considered a class of stochastic systems \( \{ X_t, t \in D \} \),

* Corresponding author. Tel.: +1-509-3352958; fax: +1-509-3358369.
E-mail address: zhanghao@wsu.edu (H. Zhang).
such that every element $Y$ in the closed linear space spanned by $X_t; t \in D$, can be associated with a series

$$L(Y) = \sum_{t \in D} a(t, Y) X_t$$

(where the sum is carried out in the order of increasing $|t|$, and $|D|$ is such that $\{t \in D, |t| \leq k\}$ is finite for any integer $k$) in such a way that

1. If the series converges in $L^2$-norm as $|t| \to \infty$, the sum equals $Y$; moreover, $L(X_t) = X_t$, for all $t \in D$.
2. For each $t$, the coefficient $a(t, Y)$ is a continuous linear functional of $Y$, i.e.

$$a(t, c_1 Y_1 + c_2 Y_2) = c_1 a(t, Y_1) + c_2 a(t, Y_2)$$

and

$$a(t, Y_n) \to a(t, Y) \quad \text{if} \quad Y_n \to Y \text{ in } L^2(dP).$$

Such a stochastic system is said to form a conditional basis. This concept was introduced for interpolation problems in the study of random fields on regular lattices and its applications have been restricted to stationary random fields on regular lattices. The definition above is a generalization of Rozanov’s to a stochastic system on an irregular lattice. We find, however, it is applicable to spatial statistics where data are collected at irregular sites. In the present work, we will study properties of a stochastic system $\{X_t; t \in D\}$ forming a conditional basis, and their applications in spatial statistics, where $D$ is either a regular or an irregular lattice in $\mathbb{R}^d$. It is organized as follows.

We will present two theorems that give necessary and sufficient conditions for a stochastic system to form a conditional basis. We will then give properties of stochastic systems forming conditional bases and discuss their applications in spatial statistics, and give an example of variogram model that is used in geostatistical data analysis, but the stochastic system does not form a conditional basis.

Let us now introduce some definition and notations. Given a stochastic system $\{X_t; t \in D\}$, denote by $H_k(X)$ the closed linear space spanned by $X_t; t \in D, |t| \geq k, k = 0, 1, 2, \ldots$, and simply write $H(X)$ for $H_0(X)$, i.e., the space spanned by all $X_t, t \in D$. For any $t$ in $D$ or not in $D$, let $\hat{X}_t$ be the projection of $X_t$ onto the space spanned by $X_s, s \in D, s \neq t$. $\{X_t; t \in D\}$ is said to be minimal if for any $t \in D$, $X_t \neq \hat{X}_t$. We assume all stochastic systems have zero means and for any $X$ and $Y$ in $L^2(dP)$, let $\langle X, Y \rangle$ and $\|X\|$ denote the inner product and norm in $L^2(dP)$, respectively.

2. Results

The following theorem was established by Rozanov (1963, Theorem 11.1, p. 105) for a stochastic system on a regular lattice. We extend it to irregular lattices. The proof follows the lines of Rozanov and is omitted here.

**Theorem 2.1.** The set $\{X_t; t \in D\}$ forms a conditional basis if and only if it is minimal and

$$\bigcap_{k \in \mathbb{N}} H_k(X) = \{0\}.$$  

In view of the theorem, if $\{X_t; t \in D \setminus D^*\}$ forms a conditional basis, then $\{X_t; t \in D\}$ also forms a conditional basis in the corresponding subspace. When $D$ is an irregular lattice, $D$ can be viewed approximately as a subset of a regular lattice that has grids fine enough to embed $D$. The next theorem gives necessary and sufficient conditions for a stationary random field on $\mathbb{Z}^d$ to form a conditional basis in terms of spectral
density, that can be verified. The proof is omitted since the theorem follows Theorem 2.1, Theorem 5.3 of Salehi and Scheidt (1972), and Theorem 4.6 of Makagon and Weron (1976).

**Theorem 2.2.** Let \( \{X_t; t \in \mathbb{Z}^d\} \) be a second order stationary random field. Then \( \{X_t; t \in \mathbb{Z}^d\} \) forms a conditional basis if and only if its spectral measure is absolutely continuous and the spectral density \( f(\lambda) \) satisfies
\[
\int \frac{1}{f(\lambda)} \, d\lambda < \infty.
\]

Properties 2.3. If \( \{X_t, t \in D\} \) forms a conditional basis, then the following holds.

(A) If \( \|X_t\| \leq M < \infty \) for any \( t \), \( \langle X_t, X_{s+t} \rangle \to 0 \) as \( |t| \to \infty \) for any \( s \in D \).

(B) \( a(t,Y) \to 0 \) as \( |t| \to \infty \) for any \( Y \in H(X) \).

(C) Let \( D_k \) consist of all points \( t \in D \) such that \( |t| \leq k \), \( k = 1,2,\ldots \). For any \( Y \in H(X) \), denote the best linear predictor of \( Y \) based on \( X_t, t \in D_k \) by
\[
Y_k = \sum_{i \in D_k} a_k(t,Y)X_i.
\]
If \( \|X_t\| \leq M < \infty \) for any \( t \), then
\[
\lim_{k \to \infty} \sup_t |a_k(t,Y) - a(t,Y)| = 0,
\]
where \( a(t,Y) \) are the coefficients in the series \( L(Y) \) associated with \( Y \).

**Proof.** We will repeatedly use the following lemma of Rozanov (1963, Lemma 2.1, p. 53): If a family of subspaces \( H_i \) of a Hilbert space \( H \) has the property that \( H_i \subset H_j \) for any \( s < t \) and \( \bigcap_{t=1}^{\infty} H_t = \{0\} \), then, for any element \( Y \in H \), its projection onto \( H_t \) converges to 0 in \( L^2 \)-norm as \( t \to \infty \).

(A) For each integer \( k \leq 1 \) and \( s \in D \), let \( X_k(s) \) be the projection of \( X_s \) onto \( H_k(X) \). Since \( \{X_s, s \in D\} \) forms a conditional basis, \( H_k(X) \) decreases to \( \{0\} \) by Theorem 2.1. Consequently, \( \|X_k(s)\| \to 0 \) as \( k \to \infty \) by Rozanov’s Lemma. Note that the projection of \( X_t \) onto \( X_{s+t} \) is
\[
\frac{\langle X_t, X_{s+t} \rangle}{\|X_{s+t}\|^2} X_{s+t},
\]
which, if \( s+t \geq k \), obviously has a norm less than or equal to \( \|X_k(s)\| \). Since \( \|X_{s+t}\| \) is bounded, \( \langle X_t, X_{s+t} \rangle \to 0 \) as \( t \to \infty \).

(B) Let \( \epsilon_t = X_t - \hat{X}_t \), \( t \in D \), we can show similarly to Rozanov (1963, Theorem 11.1, p. 105) that \( a(t,Y) = \langle Y, \epsilon_t \rangle / \|\epsilon_t\| \). The space spanned by \( \epsilon_t \) for \( t \in D \), \( |t| \geq k \), which we denote by \( H_k(\epsilon) \), is orthogonal to \( X_s \) for any \( s \) with \( |s| < k \). Consequently, \( H_k(\epsilon) \) shrinks to \( \{0\} \). Hence,
\[
\lim_{k \to \infty} \|\text{Proj}(Y|H_k(\epsilon))\| = 0.
\]
Since the projection of \( Y \) onto \( \epsilon_t \) has a norm less than or equal to \( \|\text{Proj}(Y|H_k(\epsilon))\| \) for any \( t \) such that \( |t| \geq k \), we have
\[
\frac{|\langle Y, \epsilon_t \rangle|}{\|\epsilon_t\|} \to 0 \quad \text{as} \quad |t| \to \infty.
\]

(C) By applying Rozanov’s Lemma to the orthogonal supplement of the space spanned by \( X_t, |t| \leq k \), we get \( \|Y_k - Y\| \to 0 \) as \( |t| \to \infty \). For any \( t \in D \), let \( \epsilon_t \) be defined as before. By Holder’s inequality,
\[
\sup_t |\langle Y_k, \epsilon_t \rangle - \langle Y, \epsilon_t \rangle| \leq \sup_t \|\epsilon_t\| \|Y_k - Y\| \leq M \|Y_k - Y\|.
\]
The assertion follows. \( \square \)
Remark 2.4. Property 2.3(A) provides a necessary condition for a stochastic system to form a conditional basis that can be easily verified. For example, variograms are used in geostatistics for prediction and the following variogram model has been applied to some geostatistical data (see, e.g., Cressie, 1993, p. 215),

\[
\gamma(h) = \begin{cases} 
0 & \text{if } h = 0, \\
 c_0 + c_1 |h|^\iota & \text{if } h \neq 0, \ h \in \mathbb{R}^d,
\end{cases}
\]

where \(d\) can be any integer, \(0 \leq \iota < 2\), and \(\gamma(h) = \text{Var}(X_{s+h} - X_s)/2, \forall s, h \in \mathbb{R}^d\). Note that \(\gamma(h)\) does not depend on \(s\) and such a stochastic system \(\{X_s, s \in \mathbb{R}^d\}\) is called intrinsically stationary. It is obvious that if \(\text{Var}(X_s)\) is bounded, then, for any \(s \in D, (X_s, X_{s+t}) = \text{Cov}(X_s, X_{s+t}) \to \infty \) as \(|t| \to \infty\). Therefore, \(\{X_s, s \in D\}\) cannot form a conditional basis for any countable and infinite set \(D\).

Remark 2.5. Given \(\{X_t, t \in D\}\) and an \(s \notin D\), let \(X_k(s)\) be the best predictor of \(X_s\) based on all \(X_t, t \in D_k\), and \(\hat{X}_k\) the best predictor of \(X_s\) based on all \(X_t, t \in D\). Then, following the notations in Property 2.2(C), we can write

\[
X_k(s) = \sum_{t \in D_k} a_k(t, \hat{X}_k)X_t,
\]

It is interesting to see how the coefficients \(a_k(t, \hat{X}_k)\) change when \(k\) increases. Even though it is always true that \(\|X_k(s) - \hat{X}_k\| \to 0 \) as \(k \to \infty\), the following does not always hold:

(i) \(a_k(t, \hat{X}_k)\) converges as \(k \to \infty\) for any \(t\).
(ii) \(a_k(t, \hat{X}_k)\) \(\to 0\) as both \(k\) and \(|t|\) approach \(\infty\).

Properties (B) and (C) imply (i) and (ii) hold if \(\{X_t, t \in D\}\) forms a conditional basis. An important application of (i) and (ii) is as follows.

When \(k\) is sufficiently large, the finite sum \(\sum_{t \in D_k} a(t, \hat{X}_k)X_t\) can be regarded as an approximation to \(X_k(s)\), where \(a(t, \hat{X}_k)\) are the coefficients in the series \(L(\hat{X}_k)\) associated with \(\hat{X}_k\) and, therefore, do not depend on any finite region \(D_k\). It provides an alternative method for finding the best linear predictor of \(X_s\) at unobserved site \(s\) based on a large and finite region, given the coefficients \(a(t, \hat{X}_k)\) are available. When \(\{X_t, t \in \mathbb{Z}^d\}\) is stationary whose values are known at the whole lattice except some finite points, these coefficients can be given through an algorithm (Salehi, 1979, Theorem 4). Though such an algorithm would be impossible for a random system on an irregular lattice, (i) and (ii) imply that a large and finite region will suffice for prediction in the sense that extending the region \(D\) will have little effects on the prediction error and coefficients.

Acknowledgements

We thank a referee for providing the references Makagon and Weron (1976) and Salehi and Scheidt (1972).

References