師大學報：數理與科技類
民國 89 年，45，11-19

樣本半偏差極限分佈的探討

呂小娟　　　　張少同

中國文化大學應用數學系　　國立臺灣師範大學數學系

半偏差被廣泛應用於地球統計資料的分析，而樣本半偏差便是分析者最常使用的估計量。因此很多有關於樣本半偏差的統計推論諸如半偏差模型適合度的檢定，與方向相關的對稱性質檢定等等，都須了解樣本半偏差的抽樣分佈方可進行。至今學術上曾證明得到在任一 m 相依的高斯隨機場上，某一固定位移量的樣本半偏差之邊際極限分佈。然而由於其定理條件太嚴格且由於空間性相關的因素，此結果缺乏實用性。而且只知樣本半偏差的邊際分佈仍無法進行很多統計推論。雖然有些文章曾利用模擬的方法探討樣本半偏差的聯合極限分佈，但仍欠缺理論證明。因此，在本篇文章中，我們提出一個比 m 相依更為寬鬆且為多數半偏差模型滿足的混合條件，證明在此條件下，樣本半偏差的聯合極限分佈為多變量常態。文中我們先導出在任意隨機場上，樣本半偏差的極限分佈，然後探討在高斯隨機場上，此極限分佈所須之條件及結果都大為簡化且易於應用。

關鍵詞：地球統計學、半偏差、樣本半偏差、抽樣分配

簡介

幾乎在所有地球統計學（geostatistics）的應用上，分析資料的一階及二階動差的結構為首要工作，即使整個分析的目的不在於資料結構的辨識，而是要預測隨機場在某些未取樣位置的值或判別其機率分佈的某些特性，了解資料的平均數及共變異數的性質仍然是必要的。有關空間性資料的分析，半偏差（semivariogram）被應用於各種不同的領域，諸如 Burgess and Webster（1980）及 McBratney and Webster（1986）的土壤調查（soil surveying）; Deliner and Delhome（1975）的水質分析（hydrology）; Clark（1979）及 Matheron（1971）探討採礦（mining）和礦物資源（mineral resources）的問題等等。通常地球統計資料（geostatistical data）的分析者都利用半偏差來表現資料二階變異性的特徵，所以半偏差及其估計量的統計性質已在很多學術文章上探討過，諸如 Davis and Borgman（1979）; Cressie and Hawkins（1980）; Hawkins and Cressie（1984）；以及 Robinson（1990）等文章。而最常使用的半偏差估計量莫過於樣本半偏差（sample semivariogram）。因此不論是半偏差模型適合度（goodness of fit）的檢定或任何和方向相關之對稱性質（directional symmetry properties）的檢定等應用，了解樣本半偏差的抽樣分配是極其重要的。然而至今學術上有關此抽樣分配的研究並不多，Davis and Borgman（1982）導出在任何維度的 m 相依（m-dependent）嚴格平穩隨機場（strictly stationary random fields）上，任意固定位移量（fixed lag）的樣本半偏差之極限常態分佈；Baczkowski and Mardia（1987）利用模擬的方法發現，對一個適度大小的樣本而言，樣本半偏差經對數變換後之分佈比樣本半偏差本身的分佈更近似於常態分佈。基本上，這些論文都只探討樣本半偏差在某一固定位移量的邊際分佈（marginal distribution）。然而對很多的統計推論而言，只知道樣本半偏差的邊際分佈是不足的，而且 Davis and Borgman（1982）論文中提及的 m 相依條件太嚴格，球面半偏差模型（spherical semivariogram model）是

唯一滿足此一條件的地球統計模型（geostatistical model），對於大部份經常被使用的半偏差模型則都無法適用 Davis and Borgman（1982）的結果，故其實用性受到相當大的侷限。而 Baczkowski and Mardia（1987）利用模擬方法所得的結果更須理論的探討。另一方面，至今所有學術上有關樣本半偏差抽樣分佈的文章，都侷限於高斯隨機場（Gaussian random fields）上，而未曾對其他隨機場有所探討。因此本文的主要目的是放寬 m 相依的限制至一較鬆的混合條件，並推導在任何隨機場上樣本半偏差的聯合極限分佈。由於本文使 m 相依條件大爲放寬，且大多數的半偏差模型都滿足此一條件，因此可提供大多數的地球統計資料分析者所應用。另一方面，本文亦將探討當所研究隨機場爲一高斯隨機場時，不但所需條件大爲簡化且易於檢驗，同時其樣本半偏差極限分佈的共變異矩陣（covariance matrix）更簡化爲半偏差的函數型式，使得分析者在應用上更爲簡便。

本文的結構如下：在第二節，我們先介紹和樣本半偏差抽樣分配相關之地球統計學背景，在第三節，我們分別建立樣本半偏差及其對數的聯合極限常態分佈並詳述整個推導過程，最後將結論敘述於第四節。

背景

一、平穩性和半偏差

令$\{Z(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$爲定義在 $\mathbf{R}^2$ 之子集合 $\mathbf{D}$ 上的隨機場。地球統計上的資料通常被視爲隨機場單一實現（single realization）的不完全樣本（incomplete sample），因此若沒有對隨機場一階及二階動差做任何假設，則很多統計推論是無法進行的，時間數列及空間統計學（Spatial statisitcs）上經常使用的假設爲二階平穩性（second-order stationarity），亦即假設

$$E(Z(\mathbf{s}))=\mu,\ \forall\mathbf{s}\in\mathbf{D}, \tag{1}$$

$$Cov(Z(\mathbf{s}),\ Z(\mathbf{t}))=C(\mathbf{s}-\mathbf{t}),\ \forall\mathbf{s},\mathbf{t}\in\mathbf{D}。 \tag{2}$$

(1)式中的 μ 爲一常數，(2)式中的函數 $C(\cdot)$ 稱爲共變異函數或共變異量。另一個比二階平穩性稍弱且更常爲地球統計分析者使用的性質是內部平穩性（intrinsic stationarity）。一隨機場若滿足(1)式且

$$\frac{1}{2}Var(Z(\mathbf{s})-Z(\mathbf{t}))=\gamma(\mathbf{s}-\mathbf{t}),\ \forall\mathbf{s},\mathbf{t}\in\mathbf{D}， \tag{3}$$

則稱此隨機場滿足內部平穩性。(3)式中之函數 $\gamma(\cdot)$ 即爲半偏差（semivariogram）。由半偏差的定義容易發現 $\gamma(\mathbf{h})=\gamma(-\mathbf{h})$ ，$\forall\mathbf{h}$ 且 $\gamma(\mathbf{0})=0$。一隨機場若滿足二階平穩性則必也滿足內部平穩性，且其半偏差及共變異量之關係式爲

$$\gamma(\mathbf{h})=C(\mathbf{0})-C(\mathbf{h})\ 。$$

但反之則不一定成立；亦即一內部平穩的隨機場不一定是二階平穩的隨機場，傳統的反例爲 $d-$維空間上滿足各向同性的布朗運動（isotropic Brownian motion）（參考 Robinson, 1990 and Cressie, 1993）。由於內部平穩性的適用範圍較廣及其他因素的考量（參考 Robinson, 1990），絕大部份地球統計資料分析者都利用半偏差來表示資料二階變異與位置的相關性。因此本篇文章亦將著重於半偏差的探討。

二、半偏差的估計

若一地球統計分析者得到二維空間上一內部平穩的隨機場在位置 $\mathbf{s}_1,\ldots,\mathbf{s}_n\in\mathbf{D}$ 上的觀察值。令集合$\{\mathbf{H}_1,\ldots,\mathbf{H}_k\}$爲位移量空間（lag space）$\mathbf{H}$ 的一個分割。對每一個 u=1,…,k，所有在 $\mathbf{H}_u$ 中的元素大約相等（在某種完善的定義下）且對每一個 u=1,…,k，令 $\mathbf{h}_u$ 爲集合 $\mathbf{H}_u$ 的代表值。令 $\mathbf{S}_u=\{(\mathbf{s}_i,\mathbf{s}_j):i\le j$，且 $\mathbf{s}_i-\mathbf{s}_j\in\mathbf{H}_u\}$，亦即 $\mathbf{S}_u$ 中的元素爲其距離及相對方位大約相等的位置向量，令 $\mathbf{N}_u$ 代表 $\mathbf{S}_u$ 中的元素個數。一般常用的分割爲極式分割（“polar” partition），即根據位移量的長

度及入射角分類。換言之，長度的分類$\{(0,\ r_{max}/c_L],\ (r_{max}/c_L,\ 2r_{max}/c_L],\ \ldots,\ ((c_L\text{-}1)\ r_{max}/c_L,\ r_{max}]\}$和角度的分類$\{[0,\ \pi/c_A),\ [\pi/c_A,\ 2\pi/c_A),\ \ldots,\ [(c_A\text{-}1)\pi/c_A,\ \pi)\}$（其中 r_{max} 爲兩位置間最大的距離而c_L和c_A爲指定的常數）之笛卡兒乘積（Cartesian product）即爲位移量空間的極式分割-$\mathbf{H}_1,\ldots,\mathbf{H}_v$且$v=c_Lc_A$。由 **H** 的分割則可估計上述$c_A$個方向上，每一個方位的半偏差。

假設 **H** 的一個分割已確定，則$\gamma(\mathbf{h}_u)$的估計量傳統上都利用動差法求得，亦即

$$\hat{\gamma}(\mathbf{h}_u)=\frac{1}{2\mathbf{N}_u}\sum_{(\mathbf{s}_i,\mathbf{s}_j)\in \mathbf{S}_u}\{Z(\mathbf{s}_i)\text{-}Z(\mathbf{s}_j)\}^2$$

$$(u=1,\ldots,k)。\qquad (4)$$

若對 $\mathbf{S}_u$ 中的每個元素$(\mathbf{s}_i,\mathbf{s}_j)$，$\mathbf{s}_i\text{-}\mathbf{s}_j$皆相等，(此情形可能發生，例如資料分佈在長方形的網狀格子點上）則$\hat{\gamma}(\mathbf{h}_u)$爲一不偏估計量（unbiased estimator）。於本文爲了敘述上的方便，我們將整個集合$\{\hat{\gamma}(\mathbf{h}_u):\ u=1,\ldots,k\}$稱爲樣本半偏差。

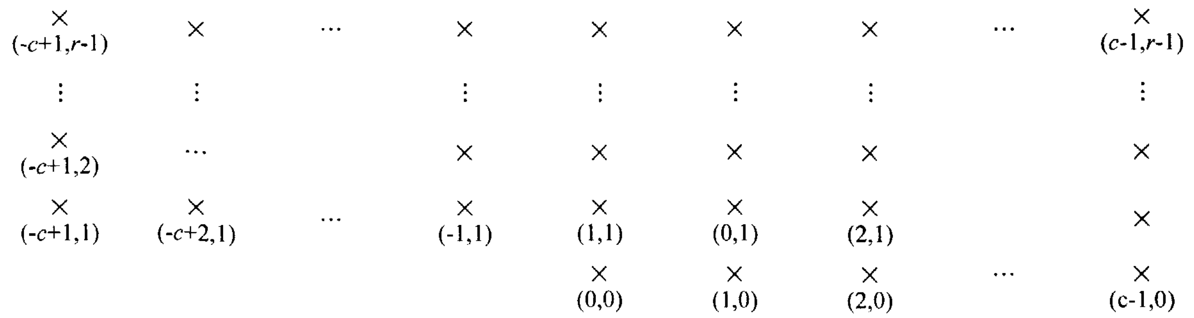
極限常態分佈

一、前言

假設$\{Z(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$爲一內部平穩的隨機場。在一大小爲$r\times c$，$x$ 和 y 方向上相鄰兩點之間隔距離固定的正規長方形格子（regular rectangular lattice）點上觀察其值，並將其觀測值記爲$\{Z(i,j):\ i=1,2,\ldots,c$ and $j=1,2,\ldots,r\}$，其中i和j分別代表行數及列數。令$\gamma(h_x,h_y)$代表 x 方向上的位移量爲 h_x 及 y 方向的位移量爲 h_y 的半偏差。在上述的格子點結構下，我們可估計集合$\mathbf{H}\equiv\{(h_x,h_y):h_x=\text{-}c+1,\ldots,c\text{-}1$ 及 $h_y=1,\ldots,r\text{-}1\}\cup\{(h_x,0):h_x=1,\ldots,c\text{-}1\}$中各個位移量的半偏差。(**H** 的結構請參考圖一）（由於相等（evenness）之故，在 **H** 中我們不需要考慮位移量 $h_y<0$ 及 $h_x<0,\ h_y=0$ 之情形)。在此情況下，樣本半偏差爲下列元素所構成：

$$\hat{\gamma}(h_x,h_y)=\frac{1}{2(c-|h_x|)(r-|h_y|)}\sum_{i=\max(1,1-h_x)}^{\min(c,c-h_x)}\sum_{j=1}^{r-h_y}(Z(i,j)-Z(i+h_x,j+h_y))^2,$$

其中 $h_x=-(c-1),-(c-2),\ldots,0,1,\ldots,c-1$ 且 $h_y=0,1,\ldots,r-1$。值得注意的是 $\hat{\gamma}(0,0)\equiv 0$，此和眞實值 $\gamma(0,0)=0$ 相吻合。另一方面，位移量很大的半偏差其估計量只能根據少數的觀測值而得，因而缺乏可靠性。一般在地球統計的實際應用上，也都只以位移量較小的估計量作爲統計推論的依據，因此以下定理將探討位移量屬於集合$\mathbf{H}'\equiv\{(h_x,h_y)\in\mathbf{H}:-l_x\le h_x\le l_x,\ 0\le h_y\le l_y,(h_x,h_y)\ne(0,0)\}$的估計量之聯合分佈，其中$0<l_x\le c-1$且$0<l_y\le r-1$。由觀察可知$\mathbf{H}'$中含有$2l_xl_y+l_x+l_y$個元素。而本文以下也將定義樣本半偏差爲所有位移量屬於$\mathbf{H}'$之半偏差的動差估計量所形成的向量。



圖一 **H** 中之位移量的結構

更明確的說，我們定義向量

$$\hat{\mathbf{G}} = [\hat{\gamma}(1,0),\ldots,\hat{\gamma}(l_x,0),\hat{\gamma}(-l_x,1),\\ \hat{\gamma}(-l_x+1,1),\ldots,\hat{\gamma}(l_x,1),\\ \hat{\gamma}(-l_x,2),\ldots,\hat{\gamma}(l_x,l_y)]'$$

爲樣本半偏差，上式之 $[\cdot]'$ 代表取轉置矩陣。$\hat{\mathbf{G}}$ 中元素的排列順序即爲圖一中由第一列開始沿著每一列由左而右的順序。令 $\mathbf{G}$ 爲相對應於 $\hat{\mathbf{G}}$ 的半偏差向量。另一方面我們亦將推導樣本半偏差經對數變換後的機率分佈 ，此原因將敘述於後。因此我們定義 $\lambda(h_x,h_y) = \log(\gamma(h_x,h_y)),\hat{\lambda}(h_x,h_y) = \log(\hat{\gamma}(h_x,h_y))$ ，且

$$\hat{\mathbf{L}} = [\hat{\lambda}(1,0),\ldots,\hat{\lambda}(l_x,0),\hat{\lambda}(-l_x,1),\\ \hat{\lambda}(-l_x+1,1),\ldots,\hat{\lambda}(l_x,1),\\ \hat{\lambda}(-l_x,2),\ldots,\hat{\lambda}(l_x,l_y)]'。$$

並令 $\mathbf{L}$ 爲相對應之眞實半偏差取對數後所形成的向量。

二、極限分佈

在推導樣本半偏差極限分佈的過程中，我們將假設方格中，行與列的數目無限增加，而兩相鄰觀察值間的單位距離保持不變。此即爲空間統計學上所謂的遞增定義域的漸近法（increasing-domain asymptotics）（參見 Cressie, 1993）。

定理 1 中，我們將推導不一定滿足常態分配性質的隨機場之樣本半偏差的極限分佈。由於證明上的須要，我們先敘述兩個性質如下。

引理 1 令 $\mathbf{X}_{(n_1,n_2)}, \mathbf{Y}_{(n_1,n_2)(j_1,j_2)}$, $n_1,n_2,j_1,j_2 \in \aleph$ 爲任意之 k 維隨機向量且滿足

(一)對任意 (j_1,j_2)，當 $n_1,n_2 \to \infty$，

$\mathbf{Y}_{(n_1,n_2)(j_1,j_2)} \xrightarrow{D} \mathbf{Y}_{(j_1,j_2)}$，

(二)當 $j_1,j_2 \to \infty$，$\mathbf{Y}_{(j_1,j_2)} \xrightarrow{D} \mathbf{Y}$，

(三)對任意 $\varepsilon > 0$，

$$\lim_{\substack{j_1\to\infty\\ j_2\to\infty}} \lim_{\substack{n_1\to\infty\\ n_2\to\infty}} \sup P(|\mathbf{X}_{(n_1,n_2)} - \mathbf{Y}_{(n_1,n_2)(j_1,j_2)}| > \varepsilon) = 0。$$

則當 $n_1,n_2 \to \infty$，$\mathbf{X}_{(n_1,n_2)} \xrightarrow{D} \mathbf{Y}$。

註：上式中 '$\xrightarrow{D}$' 代表分佈收斂。

證明：此引理爲 Brockwell and Davis（1991）書中 Proposition 6.3.9 的延伸，其證明過程極爲類似，故於此不贅述。

引理 2 令 $\{Z(\mathbf{s})\}$ 爲一二維 (m_x,m_y) 相依的嚴格平穩隨機場，並在一大小爲 $r\times c$，x,y 方向上相鄰兩點間隔距離固定的正規長方形格子點上觀察其值。假設對所有的 $\mathbf{s}$，$E\{(Z(\mathbf{s}))^4\} < \infty$，則對任意固定常數 l_x 和 l_y，當 $r,c\to\infty, \sqrt{rc}\,(\hat{\mathbf{G}}-\mathbf{G})$ 的極限分佈爲多變量常態，期望值爲 $\mathbf{0}$，共變異矩陣 Σ^{γ} 之元素爲

$$\lim_{\substack{r\to\infty\\ c\to\infty}} rcCov(\hat{\gamma}(h_x,h_y),\hat{\gamma}(g_x,g_y)) =$$

$$= \frac{1}{2}(B_{h_xh_yg_xg_y} - \frac{1}{4}B_{h_xh_yh_xh_y} - \frac{1}{4}B_{g_xg_yg_xg_y}),$$

其中 $B_{h_xh_yg_xg_y} = \frac{1}{4}\sum_{i=-m_x-b_x}^{m_x+b_x}\sum_{j=-m_y-b_y}^{m_y+b_y} C(i,j)$，$b_x = \max(|h_x|,|g_x|)$，$b_y = \max(h_y,g_y)$，且

$$C(i,j) = \\ Cov\{[(Z(t,s)-Z(t+h_x,s+h_y))^2 \\ +(Z(t,s)-Z(t+g_x,s+g_y))^2], \\ [(Z(t+i,s+j)-Z(t+i+h_x,s+j+h_y))^2 \\ +(Z(t+i,s+j)-Z(t+i+g_x,s+j+g_y))^2]\}$$

證明：此引理爲 Lu（1994）之定理 3.3.3。

定理 1 令 $\{Z(\mathbf{s})\}$ 爲一二維的嚴格平穩隨機場。在一大小爲 $r\times c$，x,y 方向上相鄰兩點間隔距離固定的正規長方形格子點上觀察 $\{Z(\mathbf{s})\}$ 的值。假設 $E\{(Z(\mathbf{s}))^4\} < \infty$ 且對任意的位移量 (h_x,h_y) 和 (g_x,g_y)，$\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}|C(i,j)| < \infty$，其中

$$C(i,j) = \\ Cov\{[(Z(t,s)-Z(t+h_x,s+h_y))^2 \\ +(Z(t,s)-Z(t+g_x,s+g_y))^2], \\ [(Z(t+i,s+j)-Z(t+i+h_x,s+j+h_y))^2 \\ +(Z(t+i,s+j)-Z(t+i+g_x,s+j+g_y))^2]\}$$

則對任意的常數 l_x 和 l_y，當 $r,c\to\infty, \sqrt{rc}\,(\hat{\mathbf{G}}-\mathbf{G})$ 的極限分佈爲多變量常態，期望值爲 $\mathbf{0}$ 共變異矩陣 Σ^{γ} 的元素爲

$$\sigma^{\gamma}_{(h_x,h_y)(g_x,g_y)}$$

$$\equiv \lim_{\substack{r\to\infty\\c\to\infty}} rcCov(\hat{\gamma}(h_x,h_y),\hat{\gamma}(g_x,g_y))$$

$$=\frac{1}{2}(B_{h_xh_yg_xg_y}-\frac{1}{4}B_{h_xh_yh_xh_y}-\frac{1}{4}B_{g_xg_yg_xg_y}),$$

其中 $B_{h_xh_yg_xg_y}=\frac{1}{4}\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}C(i,j)$

證明：令 m_x 和 m_y 爲任意二正數，定義一新隨機場 $\{Z^m(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$，使其任一元素 $Z^m(t,s)$ 滿足

(一)任一標碼集合 $\{(h_1,k_1),(h_2,k_2),...,(h_n,k_n)\}$，若對所有的 $i=1,2,...,n$，$|h_i|\le m_x$ 且 $|k_i|\le m_y$，則 $Z^m(t,s)$, $Z^m(t+h_1,s+k_1)$, $Z^m(t+h_2,s+k_2)$, ..., $Z^m(t+h_n,s+k_n)$ 的聯合分佈和 $Z(t,s)$, $Z(t+h_1,s+k_1)$, ..., $Z(t+h_n,s+k_n)$ 相同。

(二)若 $|h|>m_x$ 或 $|k|>m_y$，則 $Z^m(t,s)$ 和 $Z^m(t+h,s+k)$ 互相獨立。

如此 $\{Z^m(\mathbf{s})\}$ 爲一 (m_x,m_y) 相依的嚴格平穩隨機場且 $E\{(Z^m(\mathbf{s}))^4\}<\infty$。考慮 $\{Z^m(\mathbf{s})\}$ 上的樣本半偏差 $\hat{\mathbf{G}}^m$，亦即

$$\hat{\mathbf{G}}^m=[\hat{r}^m(1,0),...,\hat{r}^m(l_x,0),\hat{r}^m(-l_x,1),$$
$$\hat{r}^m(-l_x+1,1),...,\hat{r}^m(l_x,1),$$
$$\hat{r}^m(-l_x,2),...,\hat{r}^m(l_x,l_y)]'$$

並令 $\mathbf{G}^m$ 爲對應於 $\hat{\mathbf{G}}^m$ 之半偏差，則由引理 2，$\sqrt{rc}$ ($\mathbf{G}^m$-$\mathbf{G}^m$)的極限分佈爲多變量常態，期望値爲 $\mathbf{0}$，共變異矩陣 Σ^m 的元素爲

$$\sigma^m_{(h_x,h_y)(g_x,g_y)}$$

$$\equiv \lim_{\substack{r\to\infty\\c\to\infty}} rcCov(\hat{\gamma}^m(h_x,h_y),\hat{\gamma}^m(g_x,g_y))$$

$$=\frac{1}{2}\left(B^m_{h_xh_yg_xg_y}-\frac{1}{4}B^m_{h_xh_yh_xh_y}-\frac{1}{4}B^m_{g_xg_yg_xg_y}\right)$$

其中 $B^m_{h_xh_yg_xg_y}=\frac{1}{4}\sum_{i=-m_x-b_x}^{m_x+b_x}\sum_{j=-m_y-b_y}^{m_y+b_y}C^m(i,j)$，$b_x=max(|h_x|,|g_x|)$，$b_y=\max(h_y,g_y)$，且

$$C^m(i,j)=$$
$$Cov\{[(Z^m(t,s)-Z^m(t+h_x,s+h_y))^2$$
$$+(Z^m(t,s)-Z^m(t+g_x,s+g_y))^2],$$
$$[(Z^m(t+i,s+j)-Z^m(t+i+h_x,s+j+h_y))^2$$
$$+(Z^m(t+i,s+j)-z^m(t+i+g_x,s+j+g_y))^2]\}$$

也就是當 $r,c\to\infty$，

$$\sqrt{rc}\ (\hat{\mathbf{G}}^m\text{-}\mathbf{G}^m)\xrightarrow{D}\mathbf{Y}^m$$

其中 $\mathbf{Y}^m\sim N(\mathbf{0},\Sigma^m)$。接下來我們證明當 $m_x,m_y\to\infty$，$\Sigma^m\to\Sigma^\gamma$。爲敘述上之方便，先考慮 $h_x>g_x>0$，$h_y>g_y$ 的情形，此時

$$B^m_{h_xh_yg_xg_y}=\frac{1}{4}\sum_{i=-m_x-h_x}^{m_x+h_x}\sum_{j=-m_y-h_y}^{m_y+h_y}C^m(i,j)\text{，}$$

令

$$B^z_{h_xh_yg_xg_y}=\frac{1}{4}\sum_{i=-m_x-h_x}^{m_x+h_x}\sum_{j=-m_y-h_y}^{m_y+h_y}C(i,j)$$

則當 $m_x,m_y\to\infty$ 時，

$$B^z_{h_xh_yg_xg_y}\to B_{h_xh_yg_xg_y}\text{。}$$

考慮

$$R_{h_xh_yg_xg_y}=B^m_{h_xh_yg_xg_y}-B^z_{h_xh_yg_xg_y}\text{，}$$

則由於 $\{Z^m(\mathbf{t,s})\}$ 滿足(一)，所以當 $|i|\le m_x-h_x$ 且 $|j|\le m_y-h_y$ 時，$C^m(i,j)=C(i,j)$。又因爲 $\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}|C(i,j)|<\infty$ 及 $\{Z^m(\mathbf{t,s})\}$ 滿足(二)，所以，當 $m_x,m_y\to\infty$，

$$R_{h_xh_yg_xg_y}\to 0\text{。}\tag{5}$$

同理對於 h_x,h_y,g_x,g_y 其他情形，(5)式依然成立。亦即當 $m_x,m_y\to\infty$，

$$B^m_{h_xh_yg_xg_y}\to B_{h_xh_yg_xg_y}\text{。}$$

所以當 $m_x,m_y\to\infty$ 時，

$$\sigma^m_{(h_x,h_y)(g_x,g_y)}\to\sigma^\gamma_{(h_x,h_y)(g_x,g_y)}\text{。}$$

也就是當 $m_x,m_y\to\infty$ 時，$\Sigma^m\to\Sigma^\gamma$。因此，當 $m_x,m_y\to\infty$ 時，$\mathbf{Y}^m\xrightarrow{D}\mathbf{Y}$，其中 $\mathbf{Y}\sim N(\mathbf{0},\Sigma^\gamma)$。以下爲了利用引理 1 導出 $\sqrt{rc}$ $(\hat{\mathbf{G}}-\mathbf{G})$ 的極限分佈，我們須先證明對任意 $h=-l_x,...,l_x$，$k=0,...,l_y$ 及 $\varepsilon>0$，

$$\lim_{\substack{m_x\to\infty\\m_y\to\infty}}\lim_{\substack{r\to\infty\\c\to\infty}}\sup P\{\sqrt{rc}|\hat{\gamma}^m(h,k)-\gamma^m(h,k)$$
$$-(\hat{\gamma}(h,k)-\gamma(h,k))|>\varepsilon\}=0\text{。}\tag{6}$$

$$P\{\sqrt{rc}|\hat{\gamma}^m(h,k)-\gamma^m(h,k)$$
$$-(\hat{\gamma}(h,k)-\gamma(h,k))|>\varepsilon\}$$

$$\le\frac{rcVar(\hat{\gamma}^m(h,k)-\hat{\gamma}(h,k))}{\varepsilon^2}$$

$$= \frac{rc}{\varepsilon^2}\{Var(\hat{\gamma}^m(h,k)) + Var(\hat{\gamma}(h,k))$$

$$- 2Cov[\hat{\gamma}^m(h,k), \hat{\gamma}(h,k)]\}。$$

由於$\Sigma^m \to \Sigma^\gamma$，

$$\lim_{\substack{m_x\to\infty \\ m_y\to\infty}} \lim_{\substack{r\to\infty \\ c\to\infty}} rcVar(\hat{\gamma}^m(h,k))$$

$$= \lim_{\substack{r\to\infty \\ c\to\infty}} rcVar(\hat{\gamma}(h,k)))$$

$$= \sigma^\gamma_{(h,k)(h,k)}$$

且

$$\lim_{\substack{m_x\to\infty \\ m_y\to\infty}} \lim_{\substack{r\to\infty \\ c\to\infty}} rcCov(\hat{\gamma}^m(h,k), \hat{\gamma}(h,k))$$

$$= \sigma^\gamma_{(h,k)(h,k)}$$

因此(6)成立。而由引理 1，我們得到當$r,c\to\infty$，

$$\sqrt{rc}\ (\hat{\mathbf{G}} - \mathbf{G}) \xrightarrow{D} \mathbf{Y}。$$

至此完成定理 1 之證明。

在定理 1 的證明中，新隨機場$\{Z^m(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$的條件可稍許放寬如下：

(一)$\{Z^m(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$為一嚴格平穩隨機場。

(二)$E\{(Z^m(\mathbf{s}))^4\} < \infty$。

(三)

$$C^m(i,j)$$
$$= Cov\{[(Z^m(t,s) - Z^m(t+h_x, s+h_y))^2$$
$$+ (Z^m(t,s) - Z^m(t+g_x, s+g_y))^2],$$
$$[(Z^m(t+i,s+j) - Z^m(t+i+h_x, s+j+h_y))^2$$
$$+ (Z^m(t+i,s+j) - Z^m(t+i+g_x, s+j+g_y))^2]\}$$
$$= C(i,j)\text{，當}|h_x| < m_x, |h_y| < m_y,$$
$$|g_x| < m_x, |g_y| < m_y\text{。}$$

(四) 若 $|h| > m_x$ 或 $|k| > m_x$，則 $Z^m(t,s)$ 和 $Z^m(t+h,s+k)$，互相獨立。

如此，$\{Z^m(\mathbf{s})\}$仍為一(m_x, m_y)相依的嚴格平穩隨機場且其四次動差存在。再由上述定理 1 的推導過程，亦可證得其結果。

另一方面，為給予 Baczkowski and Mardia（1987）之模擬結果理論上的驗證，我們將建立樣本半偏差經對數變換後的極限分佈於定理 2。

定理 2 假設$\{Z(\mathbf{s})\}$為一二維隨機場且滿足定理 1 的所有條件，則當$r,c\to\infty$，$\sqrt{rc}\ (\hat{\mathbf{L}} - \mathbf{L})$的極限分佈為多變量常態，期望值為 **0**，其變異矩陣Σ^λ的元素為

$$\sigma^\lambda_{(h_x,h_y)(g_x,g_y)}$$
$$\equiv \lim_{\substack{r\to\infty \\ c\to\infty}} rcCov[\log(\hat{\gamma}(h_x,h_y)), \log(\hat{\gamma}(g_x,g_y))]$$
$$= \frac{1}{\gamma(h_x,h_y)\gamma(g_x,g_y)} \sigma^\gamma_{(h_x,h_y)(g_x,g_y)}\text{。}$$

證明：利用定理 1 的結果及多變量的 delta 方法，(參閱 Serfling, 1980, p.122)，便可得到$\sqrt{rc}\ (\hat{\mathbf{L}} - \mathbf{L})$的極限分佈。

在地球統計資料的分析上，很多分析者都假設所研究的隨機場為一高斯隨機場，實際應用上，很多地球統計資料也都符合此條件。因此，定理 3 中，我們將討論任一高斯隨機場上之樣本半偏差的極限分佈。很幸運地，我們發現在高斯的假設下，定理 1 所需的條件簡化許多且易於檢驗，而其結論中極限分佈的共變異矩陣之元素更可直接由半偏差計算而得。因此，以應用的觀點而言，定理 3 具有相當的實用價值。

定理 3 假設$\{Z(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$為一內部平穩的高斯隨機場。在一大小為$r\times c$，$x$ 和 y 方向上相鄰兩點之間隔距離固定的正規長方形格子點上觀察其值。若$\{Z(\mathbf{s})\}$之半偏差滿足

$$\sum_{q_x=-\infty}^{\infty}\sum_{q_y=-\infty}^{\infty}[\gamma(q_x+h_x, q_y+h_y)$$
$$+ \gamma(q_x-g_x, q_y-g_y) - \gamma(q_x,q_y)$$
$$- \gamma(q_x+h_x-g_x, q_y+h_y-g_y)]^2 < \infty$$

則對任意常數 l_x 和 l_y，當$r,c\to\infty$，$\sqrt{rc}\ (\hat{\mathbf{G}} - \mathbf{G})$的極限分佈為多變量常態，期望值為 **0**，共變異矩陣Σ^γ的元素為

$$\sigma^\lambda_{(h_x,h_y)(g_x,g_y)}$$
$$\equiv \lim_{\substack{r\to\infty \\ c\to\infty}} rcCov[\hat{\gamma}(h_x,h_y)), \hat{\gamma}(g_x,g_y)]$$
$$= \frac{1}{2}\sum_{q_x=-\infty}^{\infty}\sum_{q_y=-\infty}^{\infty}[\gamma(q_x+h_x, q_y+h_y)$$

$$+\gamma(q_x-g_x,q_y-g_y)-\gamma(q_x,q_y)$$
$$-\gamma(q_x+h_x-g_x,q_y+h_y-g_y)]^2$$

證明：此證明和定理 1 之推導過程非常相似。給定二常數 $m_x>0$ 、 $m_y>0$ ，定義一新隨機場 $\{Z^m(\mathbf{s}):\mathbf{s}\in\mathbf{D}\}$ 使得

$$E(Z^m(\mathbf{s}))=E(Z(\mathbf{s}))$$

且

$$C^m(h,k)$$
$$\equiv Cov[Z^m(t,s),Z^m(t+h,s+k)]$$
$$\equiv\begin{cases}C(h,k), & 若|h|\le m_x,|k|\le m_y,\\ 0, & 若|h|>m_x或|k|>m_y.\end{cases}$$

其中 $C(h,k)=Cov[Z(t,s),Z(t+h,s+k)]$ ，則 $\{Z^m(\mathbf{s})\}$ 爲一 (m_x,m_y) 相依的內部平穩高斯隨機場。令 $\hat{\mathbf{G}}^m$ 和 $\mathbf{G}^m$ 代表如同定理 1 證明中所定義之向量，由 Lu（1994）的系理 3.3.3.1，$\sqrt{rc}\ (\hat{\mathbf{G}}^m-\mathbf{G}^m)$ 的極限分佈爲多變量常態，期望値爲 $\mathbf{0}$，共變異矩陣 Σ^m 的元素爲

$$\sigma^m_{(h_x,h_y)(g_x,g_y)}$$
$$=\frac{1}{2}\sum_{q_x=-\infty}^{\infty}\sum_{q_y=-\infty}^{\infty}[\gamma^m(q_x+h_x,q_y+h_y)$$
$$+\gamma^m(q_x-g_x,q_y-g_y)-\gamma^m(q_x,q_y)$$
$$-\gamma^m(q_x+h_x-g_x,q_y+h_y-g_y)]^2$$

亦即，當 $r,c\to\infty$ ，$\sqrt{rc}\ (\hat{\mathbf{G}}^m-\mathbf{G}^m)\xrightarrow{D}\mathbf{Y}^m$，其中 $\mathbf{Y}^m\sim N(\mathbf{0},\Sigma^m)$。再利用和定理 1 相同之論證，即可得 $\sqrt{rc}\ (\hat{\mathbf{G}}-\mathbf{G})$ 的極限分佈爲多變量常態分配，期望値爲 $\mathbf{0}$，共變異矩陣 Σ^γ 的元素爲

$$\sigma^\gamma_{(h_x,h_y)(g_x,g_y)}$$
$$=\frac{1}{2}\sum_{q_x=-\infty}^{\infty}\sum_{q_y=-\infty}^{\infty}[\gamma(q_x+h_x,q_y+h_y)$$
$$+\gamma(q_x-g_x,q_y-g_y)-\gamma(q_x,q_y)$$
$$-\gamma(q_x+h_x-g_x,q_y+h_y-g_y)]^2$$

至此定成定理 3 之證明。

結論

雖然在定理 1 的結論中，樣本半偏差極限分佈的共變異矩陣非常繁雜，其條件之檢驗亦不容易，但以通用性及完整性而言，定理 1 的建立仍是必要的。且並非所有的隨機場皆滿足高斯的條件，對處理來自於非高斯隨機場之地球統計資料的分析者而言，仍可利用定理 1 求出樣本半偏差的極限分佈。因此，定理 1 仍有相當的應用價值。對高斯隨機場而言，我們在定理 3 所給的混合條件比 Davis and Bogman（1980）的 m 相依條件寬鬆許多。在各種常用的半偏差模型中，只有球面模型（spherical model）滿足 m 相依性，但我們在定理 3 所提出的混合條件爲大部分常用的半偏差模型所滿足。這些常用的模型包括許多各向同性(isotropic)或幾何非各向同性(geometrically anisotropic)的半偏差模型。例如指數模型(exponential model)，高斯模型(gaussian model)，球面模型等。且定理 3 的條件易於檢驗，而其結果亦很容易計算，故對於一般的地球統計資料分析者而言，此定理具有很高的應用價值。最後要說明的是，本文所有的研究結果可以經由類似的推導過程很容易地將其推展至任意高維度的隨機場。

參考文獻

Baczkowski, A. J. and Mardia, K. V. (1987). Approximate lognormality of the sample semivariogram under a Gaussian process. *Communication in Statistics-Simulation and Computation,* 16,571-585.

Brockwell, P. J. and Davis R. A. (1991). *Time Series: Theory and Methods, 2nd ed.,* Springer-Verlag, New York.

Burgess, T. M. and Webster, R. (1980). Optimal interpolation and isarthmic mapping of soil properties I, the semivariogram and punctual kriging. *Journal of Soil Science*, 31, 315-331.

Clark, I. (1979). *Practical Geostatistics*, Applied Sciences

Publishers, Essex, England.

Cerssie, N. (1993). *Statistics for Spatial Data.* John Wiley & Sons, New York.

Cressie, N. and Hawkins, D. M. (1980). Robust estimation of the variogram, I. *Journal of the International Association for Mathemational Geology*, 12, 115-125.

Davis, B. and Borgman, L. (1979). Some exact sampling distrbutions for varogram estimators. *Journal of the International Association for Mathernational Geology*, 11, 643-653.

Davis, B. and Brogman, L. (1982). A note on the asymptotic distribution of the sample variogram. *Journal of the International Association for Mathematical Geology*, 14, 189-193.

Delfner, P. and Delhome, J. P. (1975). Optimum interpolation by kriging. *Display and Analysis of Spatial Data* Eds. J. C. Davis and M. J. McCullogh, John Wiley & Sons, New York, 96-114.

Hawkins, D. M. and Cressie, N. (1984). Robust kriging-a proposal. *Journal of the International Association for Mathermatical Geology*, 16, 3-18.

Lu, H. C. (1994). On the asymptotic distributions of the sample covariogram and semvariogram and their use in testing for isotropy. Unpublished Ph. D. thesis, Department of Statisitics, University of Iowa.

Matheron, G. (1971). *The Theory of Regionalized Variables and its Applications.* Les Cahiers du Certre de Morphologie Mathematique, Fasc. 5, Centre de Geostatistique, Fontainebleau.

McBratney, A. B., and Webster, R. (1986). Choosing functions for semivariograms of soil properties and fitting them to sampling estimates. *Journal of Soil Science*, 37, 617-639.

Robinson, G. K. (1990). A role for variograms. *Australian Journal of Statistics*, 32, 325-335.

Serfling, R. J. (1980). *Approximation Theorem of Mathematics.* John Wiley & Sons, New York.

收稿日期： 88 年 6 月 1 日
修正日期： 89 年 4 月 10 日
接受日期： 89 年 5 月 3 日

A Note of the Asymptotic Distribution of the Sample Semivarogram

Hsiao-Chuan Lu
Department of Applied Mathematics
Chinese Culture University

Shao-Tung Chang
Department of Mathematics
National Taiwan Normal University

Abstract

The semivariogram is widely used in geostatistical data analysis. Most often, it is estimated by the sample semivariogram. Knowledge about the sampling distribuction of the sample semivariogram is necessary in many statistical inferences, such as assessing the goodness of fit of any proposed semivariogram model, and testing for directional symmetry properties. Some theoretical works have been done on the marginal distribution of the sample semivariogram for a Gaussian m-dependent process, but it is not of immediate practical use due to the restrictive conditions and the presence of spatial inter-correlations. Although the joint distrbution of the sample semivariogram has been discussed through simulation studies, theoretical support is needed. In this paper, we prove the joint asymptotic normality of the sample semivariogram under a mixing condition which is less restrictive than m-dependence and is satisfied by most geostatistical models. The distributional result is established for random fields without Gaussianity assumption and then specialized to Gaussian random fields.

Keywords: Geostatistics, Semivariogram, Sample semivariogram, Sampling distribution