

空間の統計学(6)： バリオグラムと空間的補間

慶應義塾大学総合政策学部准教授

古谷 知之 (Furutani Tomoyuki)

■兵庫県生まれ。2001年東京大学大学院工学系研究科博士課程修了。博士(工学)。東京大学大学院助手、慶應義塾大学環境情報学部専任講師を経て、07年4月より現職。専門分野：空間統計学、都市交通計画、観光政策。



1. はじめに

今回は、連続空間上でポイントサンプリングされたデータを用いて、空間的補間を行う方法を紹介します。空間的補間は、例えば、ボーリングデータを使った土壌成分や鉱石含有率の分布予測、環境観測地点データを使った環境汚染物質の分布予測等といった分野で用いられています。

空間的補間手法には、確率場データ (geostatistical data) をモデル化したバリオグラムを用いる手法と用いない方法があります。そこでまず、バリオグラムについて解説した後、空間的補間手法について紹介します。

今回の演習では、国立環境研究所「環境数値データベース」[1]の大气環境データファイルのうち、首都圏(一都三県)のSPM(浮遊粒子状物質)に関する測定局データを用いて、横浜市とその周辺のSPM分布を予測します。

Rでは、空間的補間やバリオグラム推定を行うパッケージとして、gstatやgeoRなどがあります[2][3]。本稿では、gstatを使った例を

紹介します。

```
library(spdep)
library(gstat)
spm.shp <- read.shape("tma_spm.shp")
spm <- spm.shp$att.data
ward.shp <- read.shape("Ward.shp")
ward.map <- Map2poly(ward.shp)
mesh.grid <- read.table("mesh.csv",
header=TRUE, sep=",")
coordinates(mesh.grid) <- c("X", "Y")
mesh.grid <- as(mesh.grid, "SpatialPixelsData
Frame")
```

2. バリオグラム

(1) セミバリオグラム

連続空間上の地点 $s_i (i=1, \dots, n)$ について、 s_i に関する変数を $X_k(s_i) (k=1, \dots, m)$ 、 ij 間の距離を $h_{ij} = \|s_i - s_j\|$ とします。 s_i に関する定常過程 $Z(s_i)$ が次式のように表されるとします。

$$Z(s_i) = \sum_{k=1}^m X_k(s_i) \beta_k + \varepsilon(s_i)$$

$$\varepsilon(s_i) \sim N(0, \sigma_i^2)$$

このとき、セミバリオグラム $\gamma_i(h_{ij})$ を次のように定義すると、地点間の変数の自己相関構造を表現することができます。セミバリオグラムはセミバリエンスなどとも呼ばれます。

$$\gamma_i(h_{ij}) = \frac{1}{2} E[Z(s_i) - Z(s_i + h_{ij})]^2$$

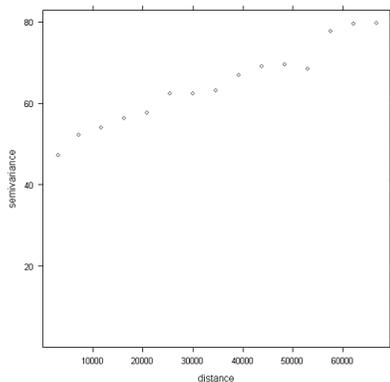
すると地点全体のセミバリオグラム $\hat{\gamma}(\tilde{h})$ は、距離 \tilde{h} となる地点 $Z(s_i)$ と $Z(s_i + h)$ の組み合わせ数 N_h により、以下のように表せます。

$$\hat{\gamma}(\tilde{h}) = \frac{1}{2N_h} \sum_{i=1}^{N_h} \gamma_i(h)$$

セミバリオグラムは、変数 $X_k(s_i)$ の設定により、様々なモデルを推定できます。例えば、 $X_k(s_i) \equiv 1$ のとき、定数項のみのモデルとなります。Rでは、**variogram()**関数を使ってバリオグラムを推定できます(図1)。

```
spm.var0 <- variogram(SPM07*1000~1,
locations=~X+Y, data=spm)
plot(spm.var0)
```

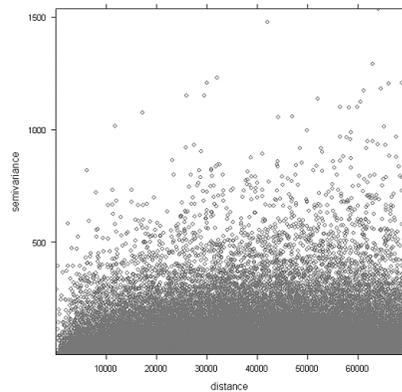
図1 バリオグラム



また変数に緯度経度のトレンドを設定することもできます。**variogram()**関数のオプション **cloud=TRUE**を指定すると、バリオグラム雲を表示できます(図2)。

```
spm.var1 <- variogram(SPM07*1000~X+Y,
locations=~X+Y, data=spm)
plot(spm.var1)
plot(variogram(SPM07*1000~X+Y,
locations=~X+Y, data=spm, cloud=TRUE))
```

図2 バリオグラム雲

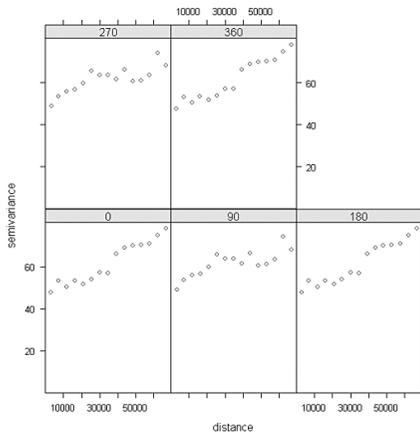


(2) 異方性バリオグラム

確率場データは、方向や方位により、自己相関の状況が変化し、バリオグラムが異なる可能性があります。全方向でバリオグラムが異なる場合、等方性 (isotropy) バリオグラムまたは全方位バリオグラムなどといいます。他方、バリオグラムを方位に応じて求めたものを、異方性 (anisotropy) バリオグラムといいます。**variogram()**関数のオプション **alpha**で方位を指定します。方位は北から順番に表示されます(図3)。

```
spm.var2 <- variogram(SPM07*1000~X+Y,
locations=~X+Y, data=spm, alpha=0:4*90)
plot(spm.var2)
```

図3 異方性バリオグラム

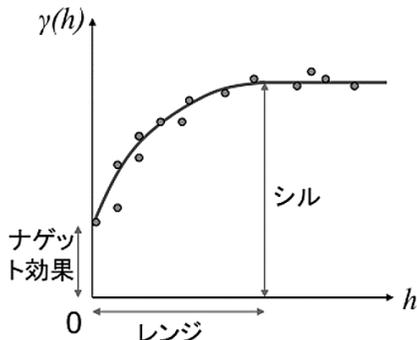


(3) バリオグラム・モデル

分布図に示されるように、バリオグラムは曲線で近似できると考えられます。しかもその曲線は定数項を持ち、一定の距離以上になるとバリオグラムが一定となる可能性があることもわかります。

一般的にバリオグラムは図4のように表されます。定数のことをナゲット効果、バリオグラムが定常状態になる距離=自己相関がなくなる距離をレンジ、データ全体の分散をシル、と呼びます。シルは、レンジに対するセミバリオグラム値でもあります。

図4 バリオグラム・モデル



代表的なバリオグラム・モデルには、次のようなものがあります。Rでは、**vgm()**関数を使ってバリオグラム・モデルを推定できます。

①指数モデル (図5)
`spm.model1 <- vgm(psill=25, model="Exp", range=28000, nugget=45)`
`plot(spm.var1, spm.model1)`

②球形モデル (図6)
`spm.model2 <- vgm(psill=25, model="Sph", range=60000, nugget=45)`
`plot(spm.var1, spm.model2)`

③線形モデル (図7)
`spm.model3 <- vgm(psill=25, model="Lin", range=56000, nugget=45)`
`plot(spm.var1, spm.model3)`

④ガウスモデル (図8)
`spm.model4 <- vgm(psill=20, model="Gau", range=35000, nugget=50)`
`plot(spm.var1, spm.model4)`

⑤ナゲット効果モデル (図9)
`spm.model5 <- vgm(psill=0, model="Nug", nugget=70)`
`plot(spm.var1, spm.model5)`

⑥Maternモデル (図10)
`spm.model6 <- vgm(psill=25, model="Mat", range=30000, nugget=45)`
`plot(spm.var1, spm.model6)`

バリオグラム・モデルに対しても異方性を考慮することができます (図11)。

`spm.anis1 <- vgm(psill=25, model="Gau", range=35000, nugget=50, anis=c(0, 0.8))`
`plot(spm.var2, spm.anis1)`

図5 指数モデル

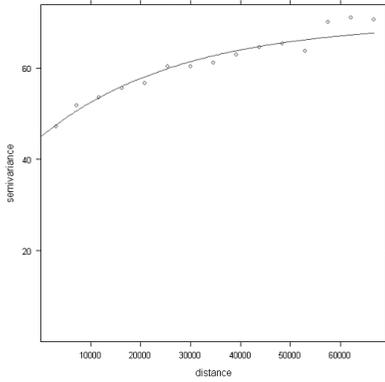


図8 ガウスモデル

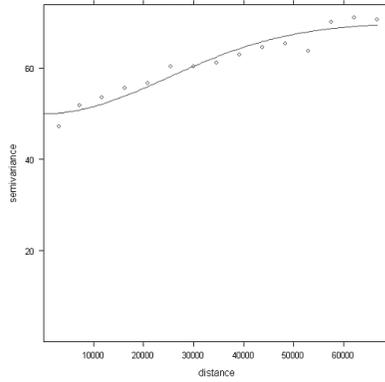


図6 球形モデル

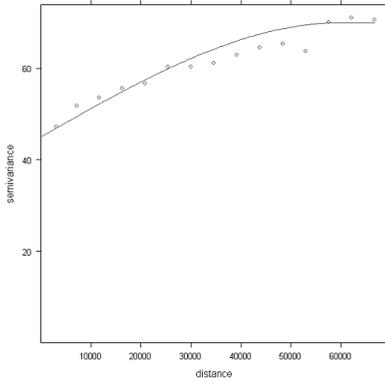


図9 ナゲット効果モデル

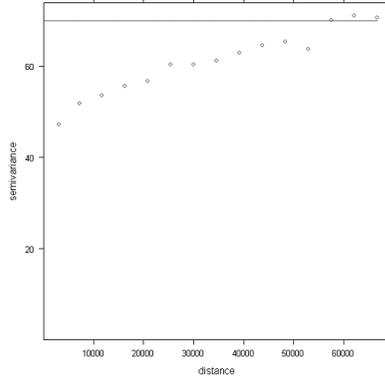


図7 線形モデル

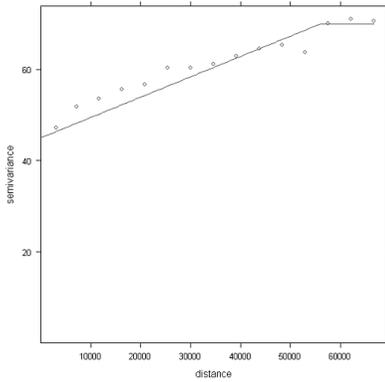


図10 Maternモデル

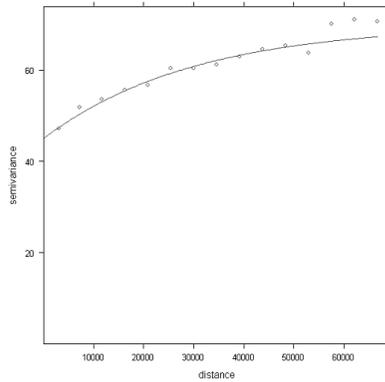


図11 異方性モデル

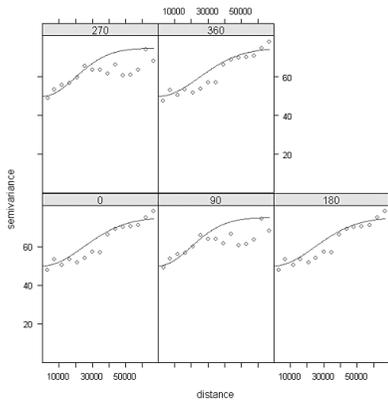
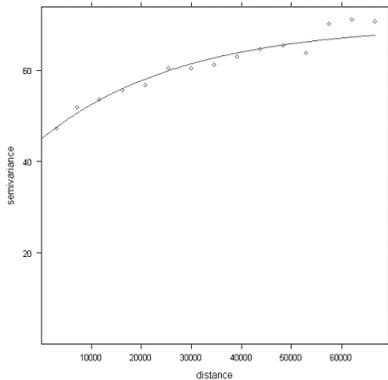


図12 バリオグラムの適合



(4) バリオグラムの適合

バリオグラム・モデルが適合しているかどうかは、`fit.variogram()` 関数を用いて示すことができます (図12)。

```
spm.fit <- fit.variogram(spm.var1, spm.model1)
plot(spm.var1, spm.model1)
```

3. 空間的補間

サンプルポイント s_i を用いて空間的補間を行う手法として、バリオグラムを用いない方法とバリオグラムを用いる方法があります。

バリオグラムを用いない方法には、逆距離加重法やスプライン近似などの手法があります。前節で紹介したバリオグラムを用いる方法には、クリギングによる補間手法とシミュレーションによる補間手法などがあります。

(1) 逆距離加重法

逆距離加重法は、地点間の距離をウェイトとする重み付け平均を求める方法です。ある地点 s_0 に対する予測値 $Z(s_0)$ は、次のように表されます。

$$Z(s_0) = \frac{\sum_{i=1}^n w(s_i) Z(s_i)}{\sum_{i=1}^n w(s_i)}$$

$$w(s_i) = \|s_0 - s_i\|^{-p}$$

Rでは、`idw()` 関数を使って逆距離加重法による空間的補間を適用できます (図13)。

```
spm.idw1 <- idw(SPM07*1000~1, locations=~X+Y, data=spm, mesh.grid, idp=2)
spplot(spm.idw1["var1.pred"], main = "ordinary kriging predictions")
```

図13 IDW

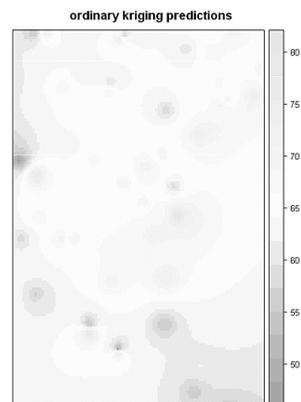


図14 単純クリギング

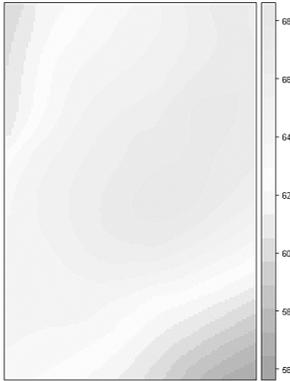


図15 通常クリギング

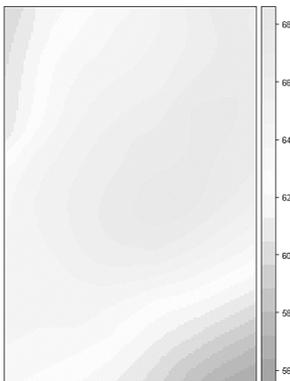
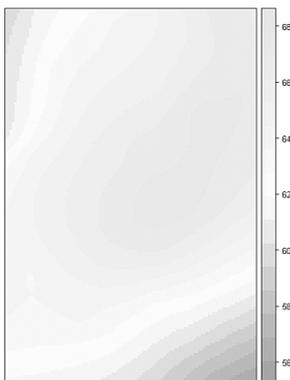


図16 普遍クリギング



(2) クリギング補間

クリギング補間は回帰モデルを用いて面的な補間を行う方法です。その主な手法として、①単純クリギング (simple kriging)、②通常クリギング (ordinary kriging)、③普遍クリギング (universal kriging) が用いられています。単純クリギングと通常クリギングはデータのトレンドを考慮せず、定数項により補間する方法です。普遍クリギングはデータの傾向面を推定する方法です。Rでは、**gstat()**関数で空間的補間を行うためのオブジェクトを生成し、**predict()**関数で空間的補間を行います。

①単純クリギング (図14)

```
spm.gs <- gstat(id="ID", formula=SPM07*
1000~1, locations=~X+Y, data=spm, model=
spm.model1, beta=mean(spm$SPM07*1000))
spm.ps <- predict(spm.gs, mesh.grid)
spplot(spm.ps[1])
```

②通常クリギング (図15)

```
spm.go <- gstat(id="ID", formula=SPM07*
1000~1, locations=~X+Y, data=spm, model=
spm.model1)
spm.po <- predict(spm.go, mesh.grid)
spplot(spm.po[1])
```

③普遍クリギング (図16)

```
spm.gu <- gstat(id="ID", formula=SPM07*
1000~X+Y, locations=~X+Y, data=spm,
model=spm.model1)
spm.pu <- predict(spm.gu, mesh.grid)
spplot(spm.pu[1])
```

* 図13～16を、SinfonicaのHP (刊行物—ESTRELA—参考) にカラーで掲載しています。

* 参考文献・URL

- [1] 国立環境研究所：環境数値データベースHP (<http://www.nies.go.jp/igreen/>).
- [2] Bivand, R.S., Pebesma, E.J. and Gomez-Rubio, V. (2008) : Applied Spatial Data Analysis with R: Springer.
- [3] Diggle, P.J. and Ribeiro Jr., P.J. (2007) : Model-based Geostatistics: Springer.