

Automatic Structure Fitting

D. Renard, N. Desassis

June 20th, 2014

Abstract

RGeostats is a library of geostatistical procedures available in the R environment. It is based on the geostatistical library **Geoslib** written in C language. The specificity of RGeostats is that it can deal with information defined in a space of any dimension, and can process any number of variables simultaneously.

This paper is meant to demonstrate the Automatic Structure Fitting procedure. It covers all the possibilities of this procedure, starting from case of a simply digitized model, to the more complex fit of a variogram map. In particular, it will cover the case of complex anisotropy, the 2-D and 3-D case.

All the demonstrations covered in this document can be reproduced as a demonstration by typing *demo(RGeostats.autofit)*.

1 Fitting a digitized variogram

This is the most simple illustration of the quality of the Automatic Model Fitting procedure.

1.1 Isotropic structure

The principle is to define an isotropic model composed of:

- a nugget effect with sill 1.2
- a spherical component with a range of 5.2 and a sill equal to 2.4

For the demonstration sake, this model (called Exdemo_ autofit1.model) is stored as an R object which belongs to the RGeostats package that can be loaded by typing the command:

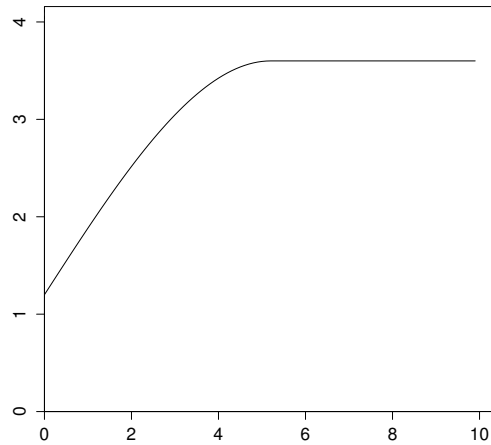
```
data(Exdemo_ autofit1.model)
```

This model is sampled by calculating an experimental variogram with 100 lags of 0.1 each. Obviously the experimental variogram exactly represents the model (with no statistical fluctuation):

```
vario <- model.sample(Exdemo_ autofit1.model,lag=0.1,nlag=100)
```

The experimental variogram is displayed in the following figure:

Figure 1: Display of the experimental sampled variogram

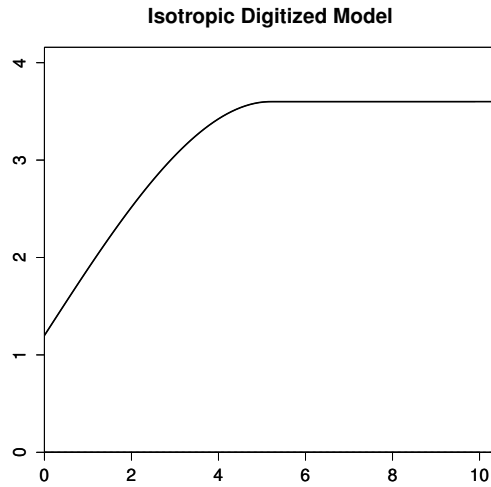


We now wish to fit a model to this experimental variogram using the default parameters, i.e. a mixture of basic structures such as Nugget Effect, Gaussian, Cubic, Exponential and Spherical components:

```
model <- model.auto(vario)
```

The previous command automatically fits the optimal model (which will be stored in the object called model, and plot the model together with the experimental variogram, as shown in the next figure. As the fit is perfect, we cannot distinguish the two curves:

Figure 2: Display of the experimental variogram and the fitted model



When comparing the characteristics of the fitted model to the ones of the digitized one, we can check that the procedure has performed perfectly.

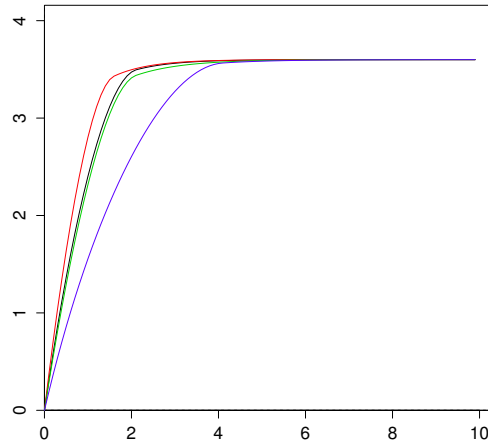
1.2 Anisotropic structure

The same work-flow is performed on an model in the 2-D space with nested structures and complex anisotropies:

- a basic exponential structure with a sill of 1.2 and ranges varying from 2.4 to 3.6 and an anisotropy rotation angle of 20 degrees
- a basic spherical component with a sill of of 2.4 and ranges varying from 4.1 to 1.6 and an anisotropy rotation angle of -45 degrees

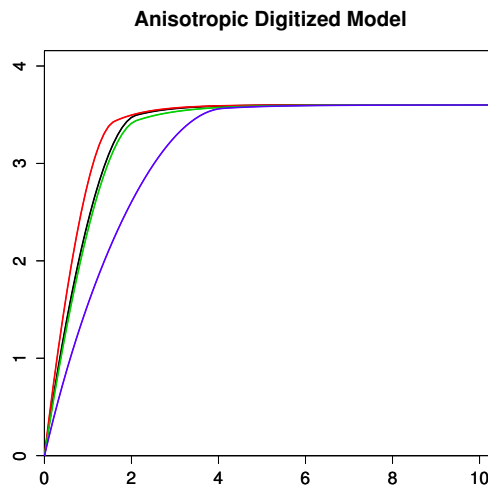
The model is sampled in the 4 main directions (0, 45, 90 and 135 degrees) and represented in the next figure:

Figure 3: Display of the experimental sampled anisotropic variogram



After 1000 iterations, a reasonably good fit is obtained as can be checked on the following figure:

Figure 4: Display of the experimental variogram and the corresponding model



Note that the characteristics of the model are found by the Automatic Model Fitting even if the four directions along which the experimental variograms have been calculated do not match main orientation of the anisotropy ellipsoid.

2 Fitting a real experimental variogram

In this second part, we perform the Automatic Structure Fitting from real experimental variograms.

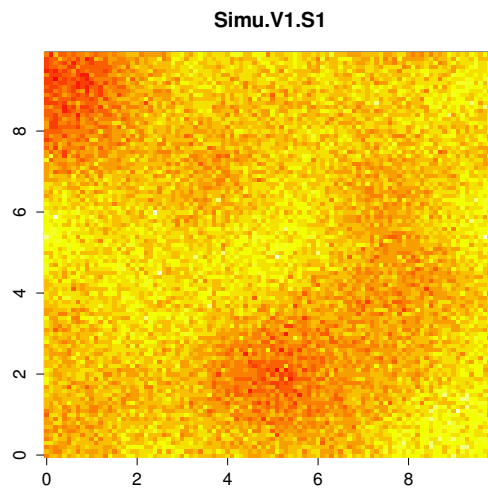
2.1 The 2-D case

We define a 2-D model first composed of:

- a nugget effect with sill equal to 1.2
- an gaussian component with a sill equal to 1.5 and ranges varying from 3.5 to 2.6 and an anisotropy rotation angle of 20 degrees.

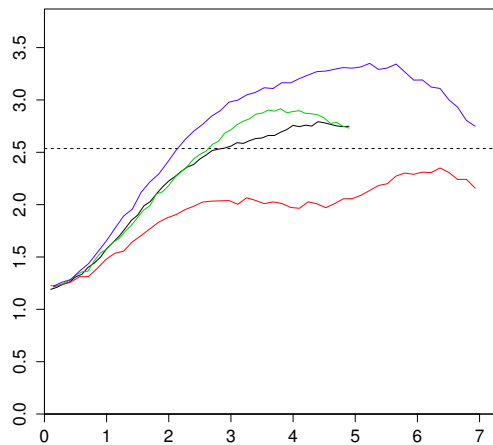
We calculate one simulation outcome (using the turning band method) of the model on a fine 2-D square grid with 100 nodes of 0.1 along each axis. The outcome is displayed in the following figure.

Figure 5: Simulation outcome



The experimental variograms calculated from this grid in the four main directions and stored in the R object `Autof1.vario`. The experimental variogram is displayed in the next figure:

Figure 6: Experimental variogram calculated on the simulation outcome in 4 directions

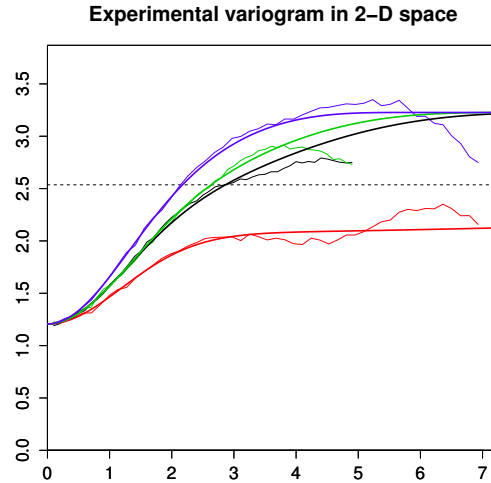


For demonstration sake, this R object has been saved in the package and can be loaded by typing:

```
data(Exdemo_autofit3.vario)
```

This variogram (computed in 4 directions) is fitted using the Automatic Model Fitting procedure. The resulting model is displayed together with the experimental variograms in the next figure:

Figure 7: Experimental variogram and the corresponding model



The fit is graphically acceptable, even if its characteristics are quite different from the ones of the initial model:

- nugget effect with a sill equal to 1.206
- a gaussian component with a sill equal to 0.863, a range varying between 2.773 and 2.969 and an anisotropy rotation angle of -11 degrees
- a cubic component with a sill equal to 5.929, a zonal anisotropy in the direction -50 degrees and a range equal to 1.157

2.2 The 3-D case

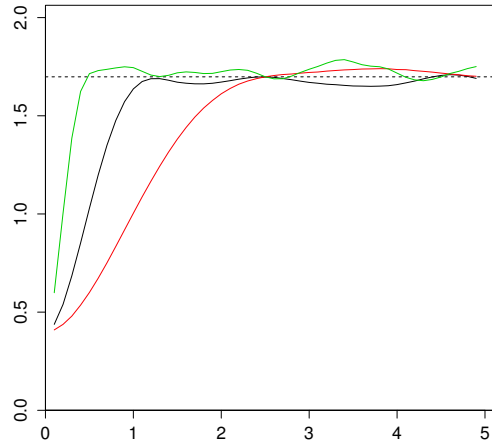
We perform the same type of operation in the 3-D space.

The model is composed of:

- a nugget effect with a sill equal to 0.4
- a cubic component with a sill equal to 1.3 and ranges equal to 1.5 along X, 3 along Y and 0.6 along Z.

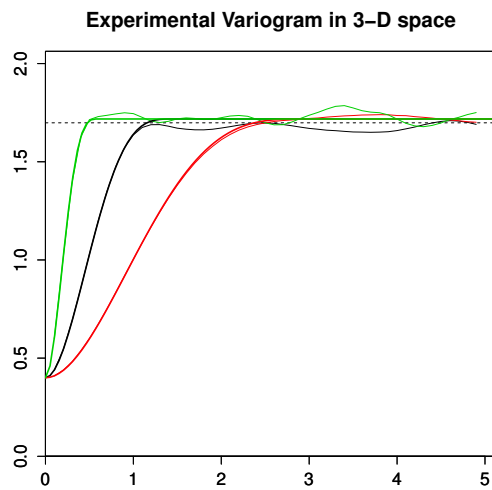
As we can see, we do not let the automatic model fitting procedure find the angles, which reduces the number of parameters to be fitted and therefore reduce the consuming time. The experimental variogram is calculated along the three main axes of the grid, with 50 lags, as represented in the following figure:

Figure 8: Experimental variogram along the main directions in the 3-D space



After 1000 iterations, the model perfectly reproduces the experimental variogram as shown in the following figure:

Figure 9: Experimental variogram and the corresponding model



The characteristics of the fitted model are:

- a nugget effect with a sill equal to 0.4

- a cubic component with a sill equal to 1.318 and ranges equal to 1.481 along X, 3.047 along Y and 0.609 along Z.

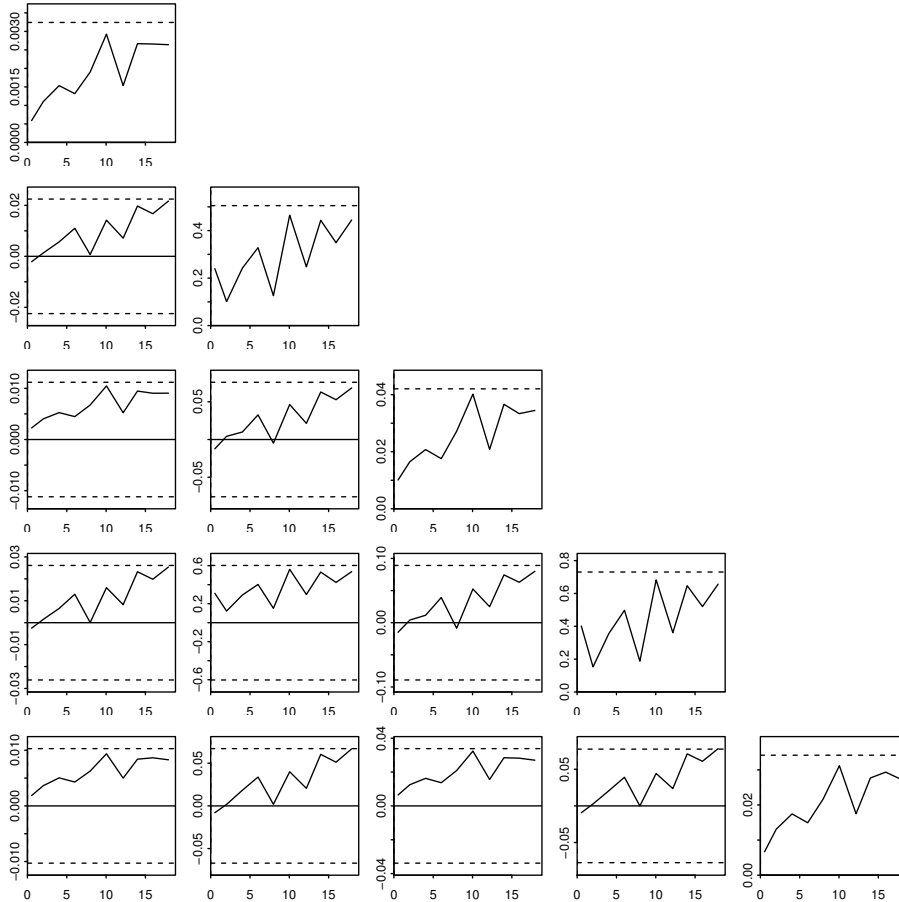
3 Fitting in the multivariate case

This example deals with a 2-D case with a large number of variables (18). The experimental variogram already has been calculated and stored as an R object that can be downloaded by simply typing:

```
data(Exdemo_ Exdemo_ autofit5.vario)
```

The interested user can visualize the experimental variogram. However, due to the large number of variables, there would be $18 \cdot 19 / 2 = 171$ simple and cross-variograms, each one represented on a very small portion of the figure. This is the reason why, by default, only the simple and cross-variograms of the first five variables are visualized.

Figure 10: Multivariate Experimental variogram

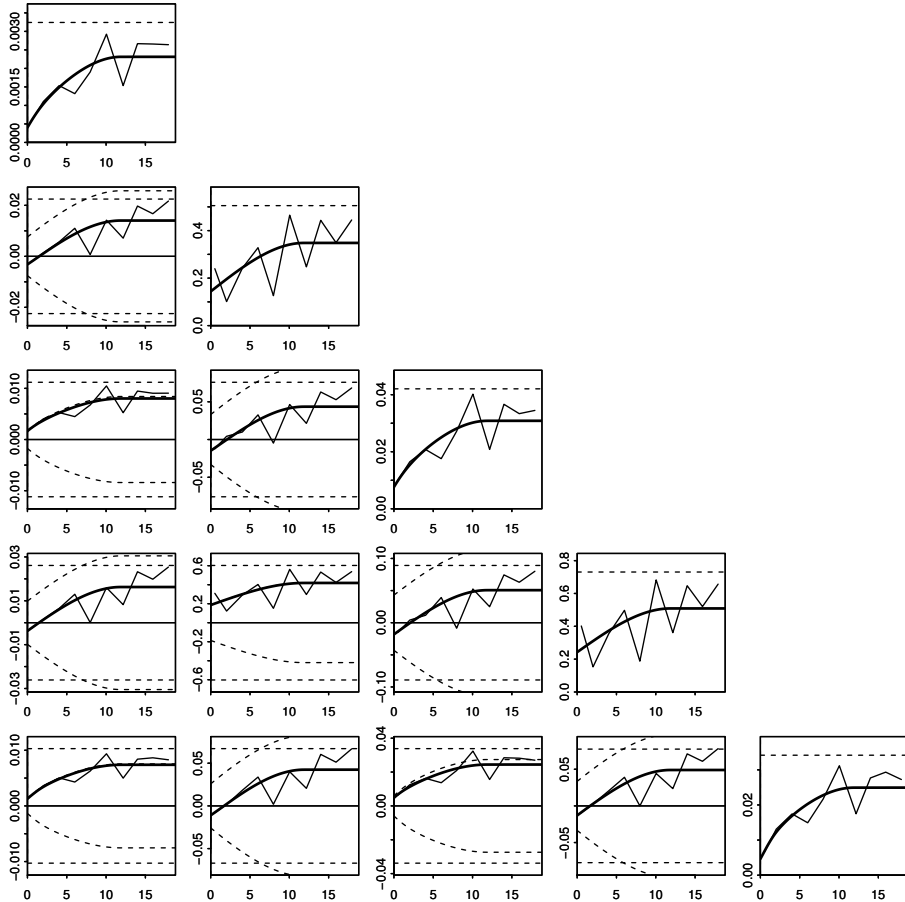


The automatic model fitting procedure is performed using the defaulted five basic structures (Nugget Effect, Gaussian, Cubic, Exponential and Spherical components) and having the Goulard option switched on.

```
model <- model.auto(Exdemo_autofit5.vario)
```

As there is only one calculation direction, we are not looking for any anisotropy: each model is characterized by its range (except the Nugget Effect) and the matrix of sills (of dimension 18*18). Using the Goulard option reduces the number of parameters to be fitted within the Foxleg algorithm down to 4 parameters (the ranges of each structure). All five basic structures are kept in the final model which is represented in the next figure:

Figure 11: Multivariate Experimental variogram and the corresponding model



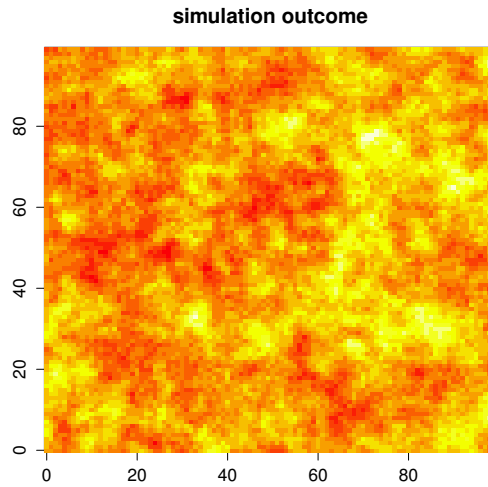
4 Breaking the ambiguity

This example has been requested by a reviewer of the paper presenting the method. We use a regular 2-D square grid with 100 nodes and a mesh of 1m: the grid covers 100m by 100m. We simulate the outcome of a random function which follows a structure composed of three nested isotropic spherical components:

- the small range component (5m) with a sill of 1.5
- the middle range component (12m) with a sill of 2
- the long range component (28m) with a sill of 1.4

The simulation outcome is represented in the next figure:

Figure 12: Simulation Outcome

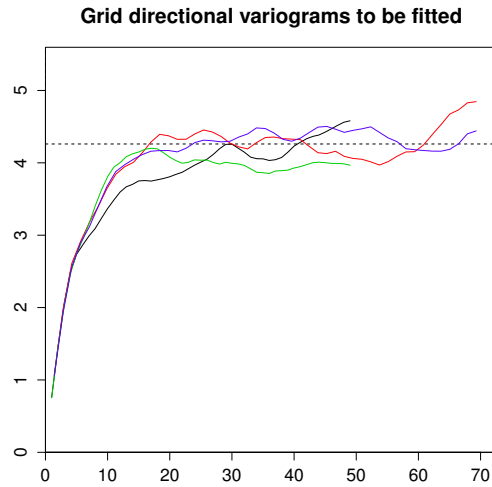


The variogram is calculated along the four main directions (two axes and two bisectors) of the grid using the following statement:

```
dirincr <- matrix(c(1,0,1,1,0,1,-1,1),nrow=2,ncol=4)
vario <- vario.grid(grid,dirincr,nlag=50)
```

which gives the following graphic representation. Note that, although the model is isotropic, the different experimental variograms shows some statistical fluctuations that could be interpreted as anisotropy.

Figure 13: Experimental directional variograms on the grid



This experimental variogram is saved as an R object that can be downloaded directly by typing:

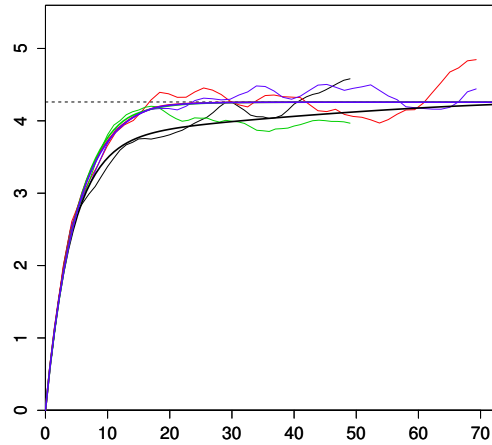
```
data(Exdemo_ autofit6.vario)
```

The first attempt consists in running the automatic model fitting procedure using the defaulted five basic structures (Nugget Effect, Gaussian, Cubic, Exponential and Spherical components):

```
model <- model.auto(Exdemo_ autofit6.vario)
```

leading to the following result composed of anisotropic exponential and spherical components. The fit is reasonably correct although the convergence has not been reached after 1000 iterations.

Figure 14: Model with no constraint



The next trial is performed by defining the set of possible basic structures as a combination of three spherical components, as follows:

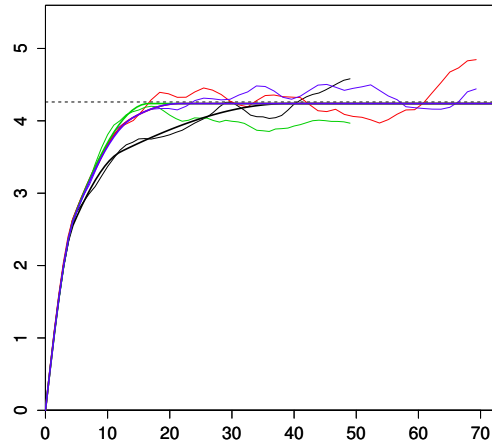
```
model <- model.auto(Exdemo_ autofit6.vario,struct=melem.name(c(3,3,3)))
```

As expected, the resulting model is composed of three nested anisotropic spherical structures:

- a spherical component with range varying from 1.61 to 1.71 and a sill of 4.437
- a spherical component with range varying from 12.11 to 16.61 and a sill of 1.419
- a spherical component with range varying from 16.58 to 39.04 and a sill of 1.211

The four directions of the model are displayed in the following figure:

Figure 15: Model with three nested spherical components



5 Fitting a Variogram Map

The principle is to create a square grid with 200 nodes and a mesh of 1m:

```
grid <- db.create(flag.grid=T,nx=c(200,200))
```

The model is stored as an R object and can be loaded using the statement:

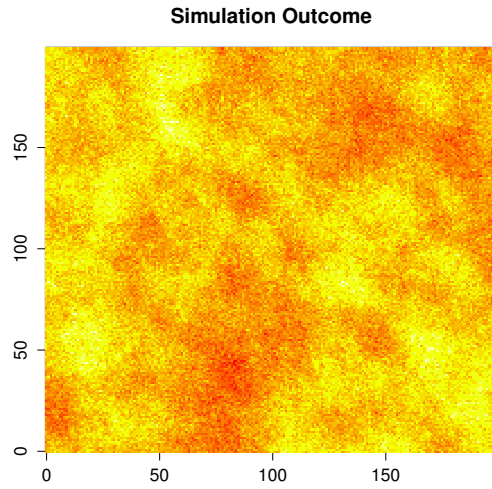
```
data(Exdemo_autofit7.model)
```

A simulation outcome is calculated over the grid using the Turning Bands simulation technique (with 1000 bands):

```
grid <- simtub(,grid,Exdemo_autofit7.model,nbtuba=1000)
```

Plotting the simulated outcome produces the following figure:

Figure 16: Simulation outcome

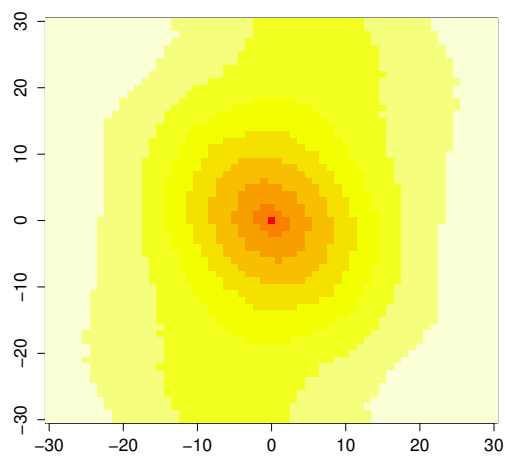


The next step consists in calculating the variogram map on the previous outcome: the extension of the variogram map is fixed to 61 by 61 nodes (i.e. the radius is equal to 30 nodes).

```
vmap <- vmap.grid(grid,nx=30,ny=30)
```

which leads to the following variogram map representation:

Figure 17: Variogram map



The Automatic Variogram Map Fitting procedure is performed by defining the set of possible basic structures to a combination of a nugget effect and two spherical components:

```
model <- vmap.auto(vmap,struct=melem.name(c(1,3,3)))
```

Although the convergence has not been reached after 1000 iterations, the graphic result is quite convincing, where the initial experimental variogram map is displayed together with the theoretical variogram map obtained with the fitted model:

