Geostat RDV - Perth - February 2013

Automatic variogram modeling

N. Desassis, D. Renard

Geovariances

Where no one has gone before

Content

- Introduction
- How does it work ?
- Illustration (variogram map)
- Multivariate case

That old chestnut

Why an automatic fitting ?

- Need for a model
  - ✓ Variogram, variogram map, cross-variograms,...
- High number of parameters
  - ✓ 3 D, anisotropies, multivariate, nested structures...
- A lot of models required
  - ✓ Local Geostatistics
- Indirect observations of the quantity to fit
  - ✓ PGS, regularisation, ...

Specifications

The algorihm has to be

- Automatic
- Fast
- Robust

How does it usually work ?

- Black box = computer program

How does it usually work ?

Inputs

- n experimental variograms $\hat{\gamma}(\vec{h}_j)$

Geovariances

How does it usually work ?

Inputs

- n experimental variograms $\hat{\gamma}(\vec{h}_j)$

Output

- A fitted model $\gamma(\vec{h})$

Geovariances

How does it usually work ?

Inputs

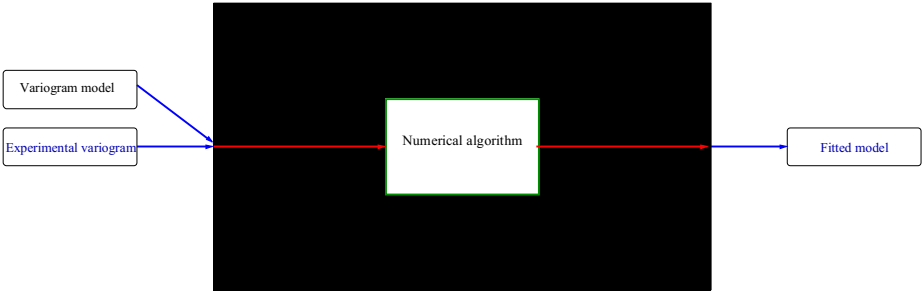
- n experimental variograms $\hat{\gamma}(\vec{h_j})$
- a model $\gamma_\psi(\vec{h})$ with a set of unknown parameters ψ

Output

- the "best" parameters $\psi^\star$

Geovariances

How does it usually work ?

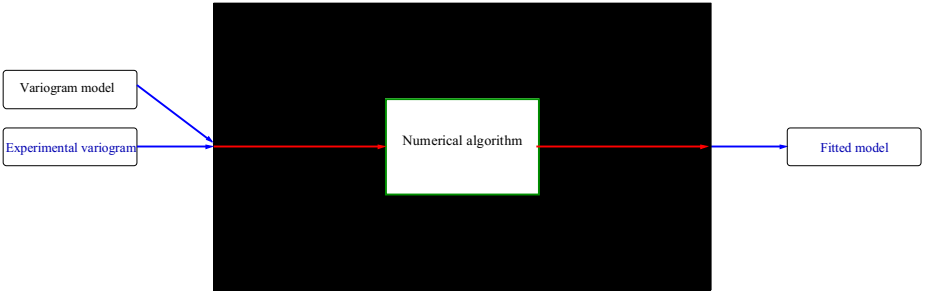


The "best" parameters $\psi^\star$ is the minimum of the cost function

$$S(\psi) = \frac{1}{2} \sum_{j=1}^{n} \omega_j r_j(\psi)^2$$

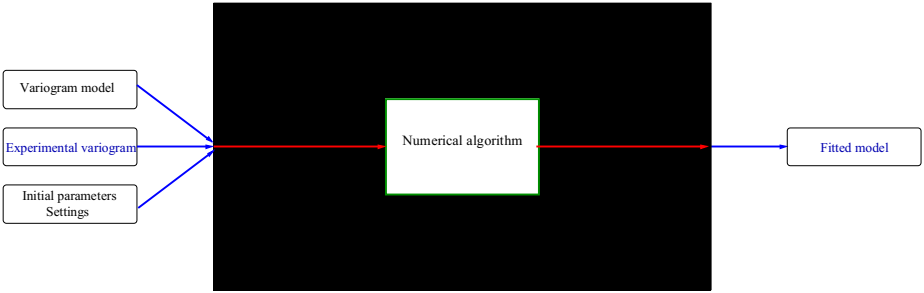
with residuals $r_j(\psi) = \gamma_\psi(\vec{h_j}) - \hat{\gamma}(\vec{h_j})$

How does it usually work ?



- Iterative algorithm
- It produces a sequence $\psi_1, \psi_2, ...$
- It converges to the minimum $\psi^\star$ of the cost function $S(\psi)$

How does it usually work ?



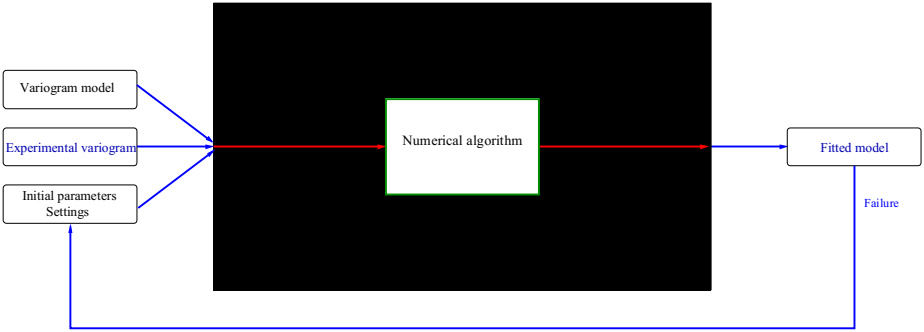
Inputs

- n experimental variograms $\hat{\gamma}(\vec{h_j})$
- a model $\gamma_\psi(\vec{h})$ with a set of unknown parameters ψ
- require ψ_0 and other algorithm settings

Output

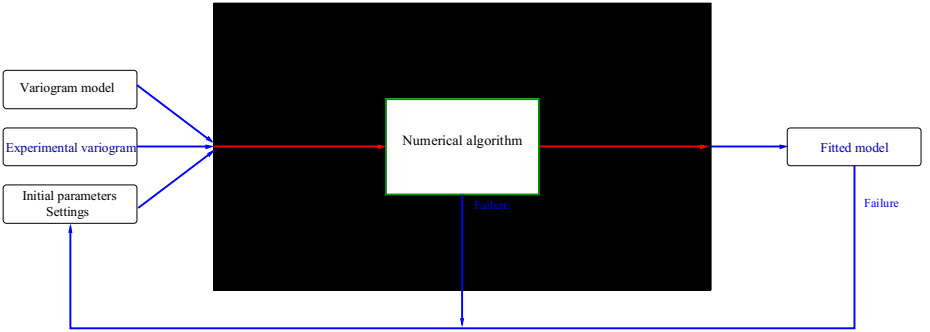
- the "best" parameters $\psi^\star$

How does it usually work ?



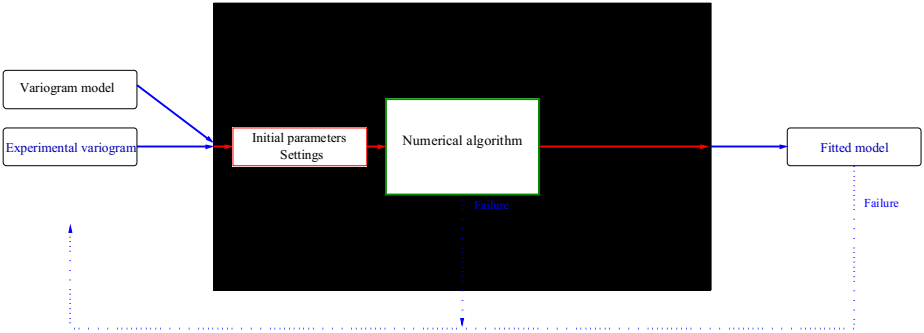
- The result is often not satisfactorily...

How does it usually work ?



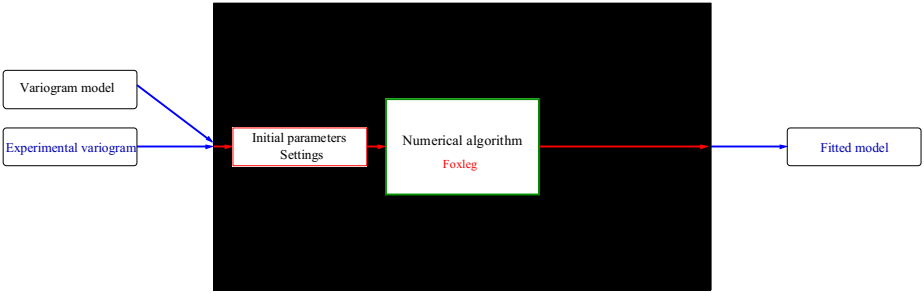
- The result is often not satisfactorily...
- ... and sometimes the algorithm failed (error)

How does it work in Isatis© ?



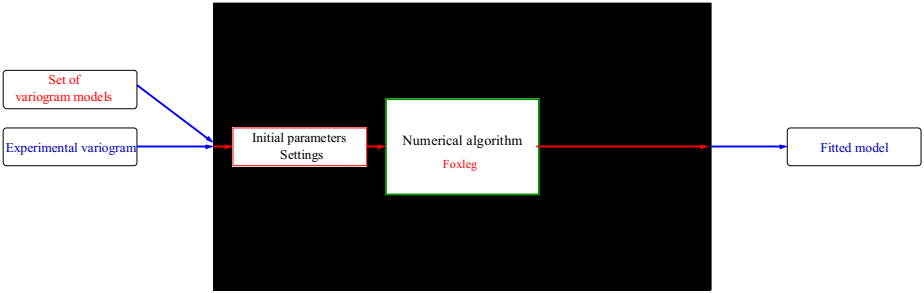
- Good settings and starting points are automatically choosen

How does it work in Isatis© ?



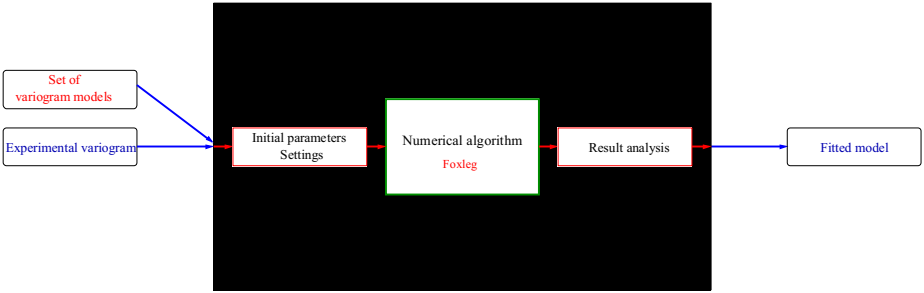
- Good settings and starting points are automatically choosen
- The numerical algorithm is adapted to avoid problems

How does it work in Isatis© ?



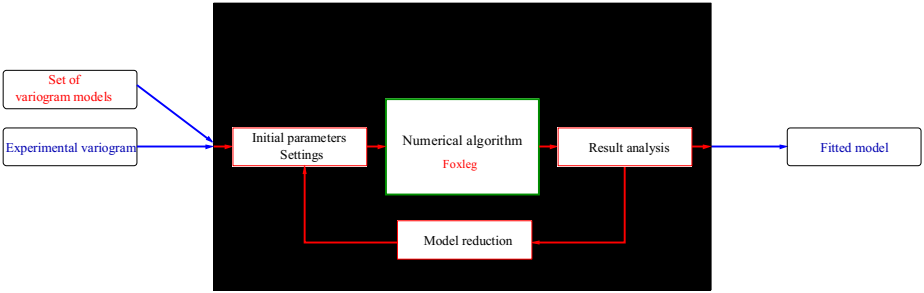
- A family of elementary models with unit sills $g^{(1)}, \ldots, g^{(p)}$
- Parameters θ_i : $g^{(i)} \equiv g^{(i)}_{\theta_i}$.
- Find a model $\gamma_\psi = \sum_{i=1}^{p} \lambda_i g^{(i)}_{\theta_i}$

How does it work in Isatis© ?



- analysis of the results
- percentage of variances explained by each basic structure

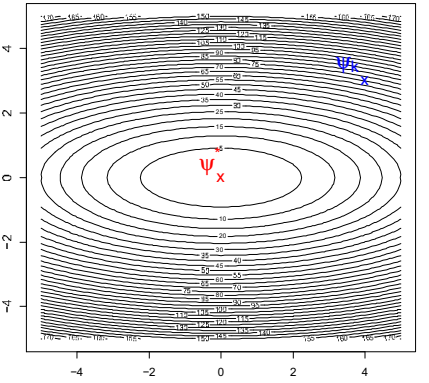
How does it work in Isatis© ?



- analysis of the results
- percentage of variances explained by each basic structure
- supression if below a threshold

Descent methods

How to find $\psi^\star$?



Geovariances

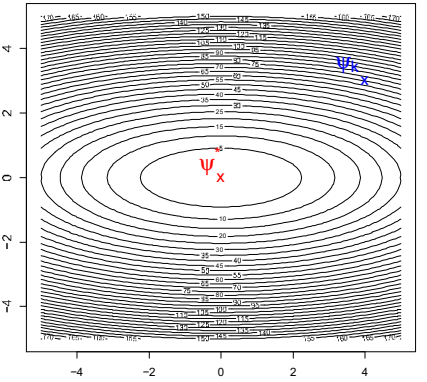
Descent methods

How to find $\psi^\star$?

- Consider $\alpha > 0$ and ε an unitary vector
- Taylor expansion around ψ_k along the direction ε

$$S(\psi_k + \alpha\varepsilon) \simeq S(\psi_k) + \alpha\varepsilon^t \nabla S(\psi_k)$$

- ε is a descent direction if $\varepsilon^t \nabla S(\psi_k) < 0$



Gecvariances

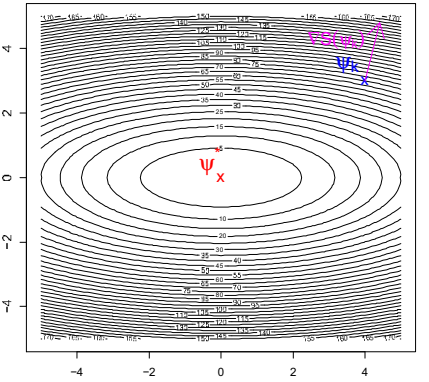
Descent methods

How to find $\psi^\star$?

- Consider $\alpha > 0$ and ε an unitary vector
- Taylor expansion around ψ_k along the direction ε

$$S(\psi_k + \alpha\varepsilon) \simeq S(\psi_k) + \alpha\varepsilon^t \nabla S(\psi_k)$$

- ε is a descent direction if $\varepsilon^t \nabla S(\psi_k) < 0$



Gecvariances

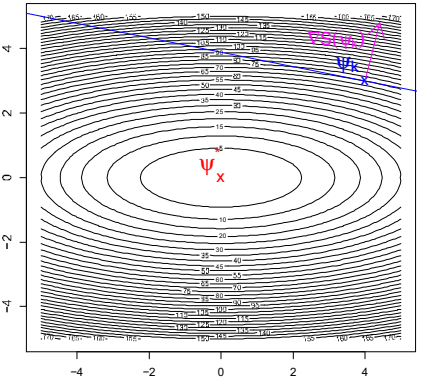
Gradient method

How to find $\psi^\star$?

- Consider $\alpha > 0$ and ε an unitary vector
- Taylor expansion around ψ_k along the direction ε

$$S(\psi_k + \alpha\varepsilon) \simeq S(\psi_k) + \alpha\varepsilon^t \nabla S(\psi_k)$$

- ε is a descent direction if $\varepsilon^t \nabla S(\psi_k) < 0$



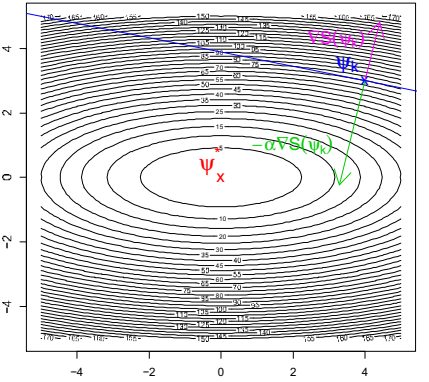
Gecvariances

Gradient method

- Relative gain for $\alpha > 0$:

$$\lim_{\alpha \to 0} \frac{S(\psi_k) - S(\psi_k + \alpha\varepsilon)}{\alpha||\varepsilon||} = -\frac{1}{||\varepsilon||}\varepsilon^t \nabla S(\psi_k) = -||\nabla S(\psi_k)|| cos(\phi)$$

- $\varepsilon = -\nabla S(\psi_k)$ locally the best direction
- $\psi_{k+1} = \psi_k - \alpha \nabla S(\psi_k)$
- need to find α
- succesive linear approximation of S
- Robust and well adapted far from the minimum (first iterations)
- Very slow convergence



Newton-Raphson method

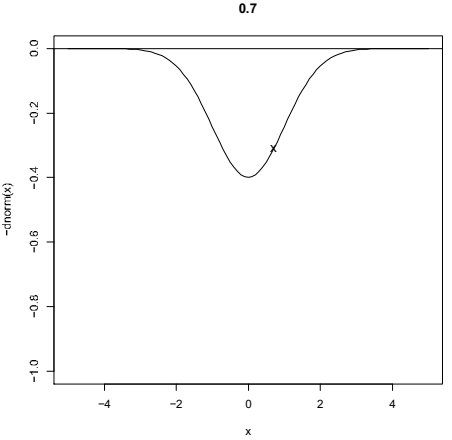
- Second order Taylor expansion (quadratic approximation of S)

$$S(\psi_k + \varepsilon) \simeq L(\varepsilon) = S(\psi_k) + \varepsilon^t \nabla S(\psi_k) + \frac{1}{2}\varepsilon^t H_S(\psi_k)\varepsilon$$

- $H_S(\psi_k)$ matrice of the second order derivatives of S computed for ψ_k
- Minimize $L(\varepsilon)$ by equating L to 0
- It leads to

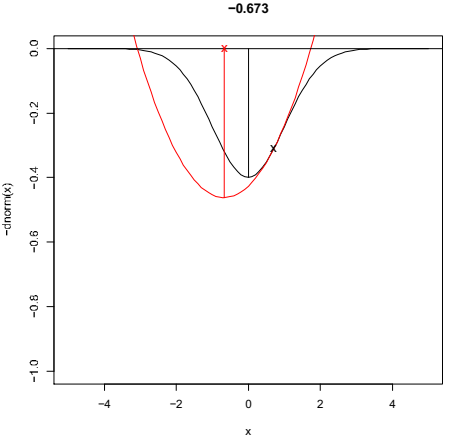
$$\psi_{k+1} = \psi_k - H_S(\psi_k)^{-1} \nabla S(\psi_k)$$

Exemple with a good case



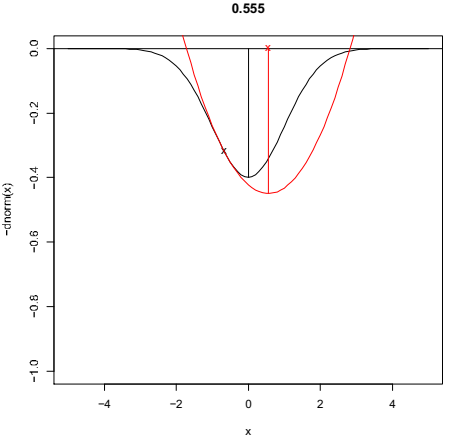
Gecvariances

Exemple with a good case



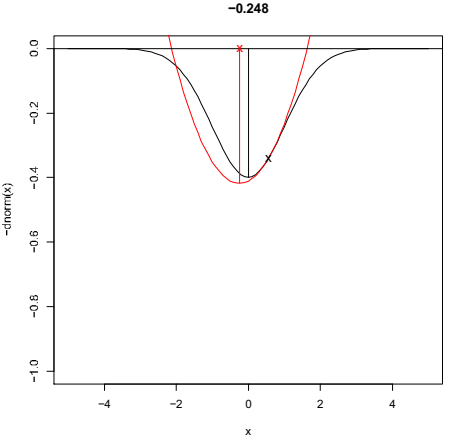
Gecvariances

Exemple with a good case



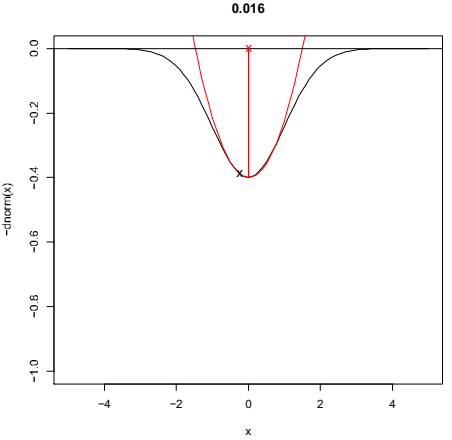
Gecvariances

Exemple with a good case



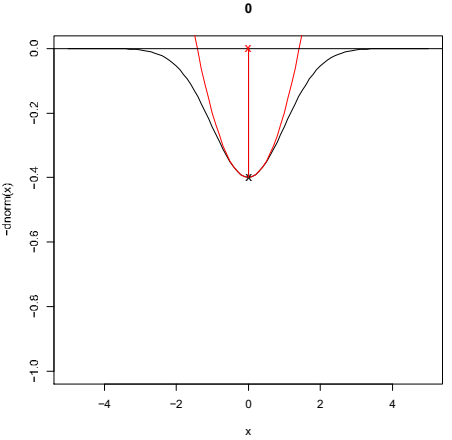
Gecvariances

Exemple with a good case



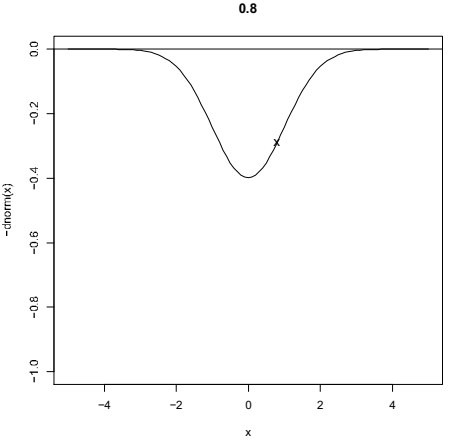
Gecvariances

Exemple with a good case



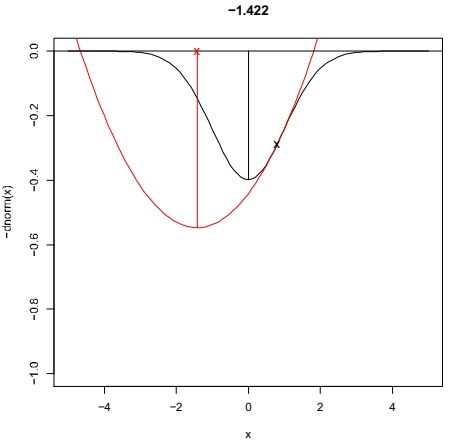
Gecvariances

Exemple with a bad case



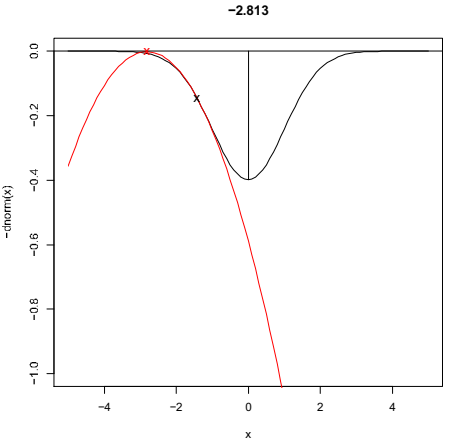
Gecvariances

Exemple with a bad case



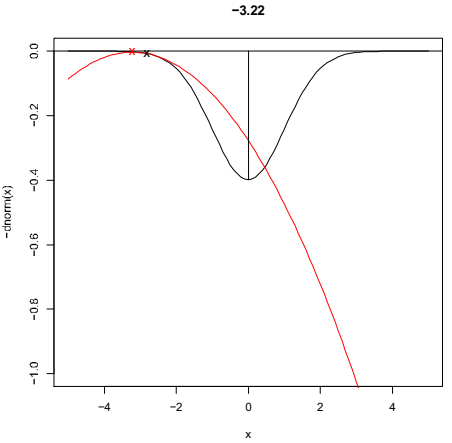
Gecvariances

Exemple with a bad case



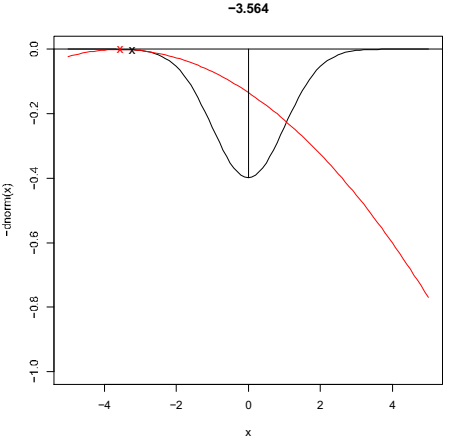
Gecvariances

Exemple with a bad case



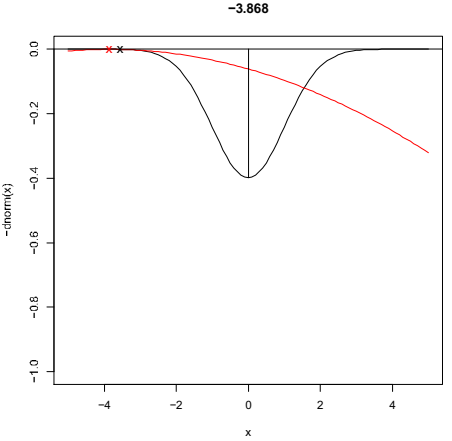
Gecvariances

Exemple with a bad case



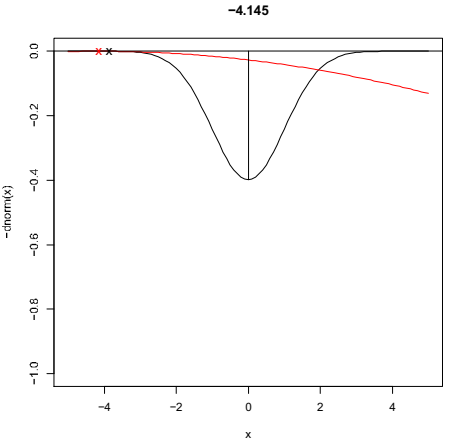
Gecvariances

Exemple with a bad case



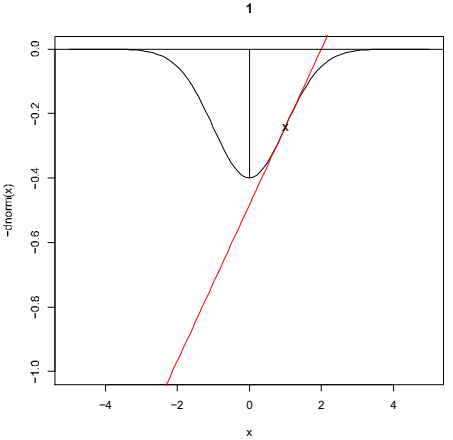
Gecvariances

Exemple with a bad case



Gecvariances

Singularity



Properties of Newton-Raphson

- Great convergence speed not far from the minimum
- Lack of robustness far from the minimum
- Problem when $H_S(\psi)$ is not invertible

Singularities of $H_S(\psi)$

- sub-space of minima
- replace the inverse of $H_S(\psi)$ (which doesn't exist) with a generalized inverse to find a particular minimum of L
- Moore-Penrose inverse (Singular Values Decomposition)

Trust region

- Find a compromise between robustness and convergence speed
- Idea : we suppose that the quadratic approximation is only correct inside a ball around ψ_k with radius Δ_k
- We are looking for the minimum of L under the constraint that $||\varepsilon||_1 \leq \Delta$
- We check the quality of the approximation

$$r = \frac{S(\psi_{k+1}) - S(\psi_k)}{L(\psi_{k+1}) - L(\mathbf{0})}$$

- If S has decreased and if the approximation has correctly predicted the observed diminution of S (r closed to 1) we increase the trust radius Δ_{k+1} for the next iteration
- Otherwise it is decreased
- Furtheremore, ψ_{k+1} is not accepted if S increases

Gecvariances

Illustration 1

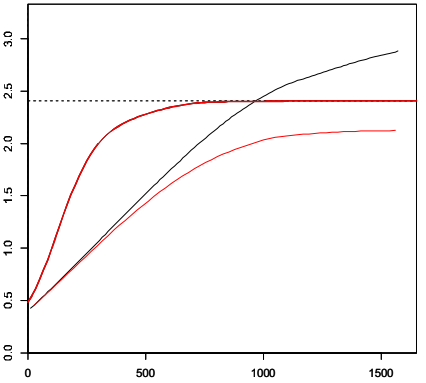
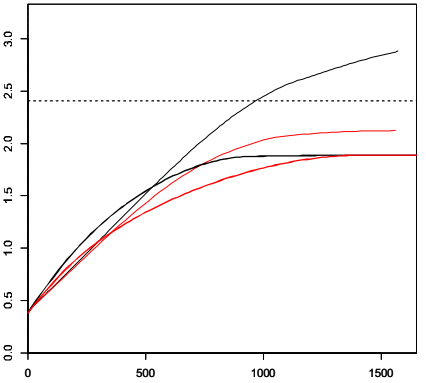
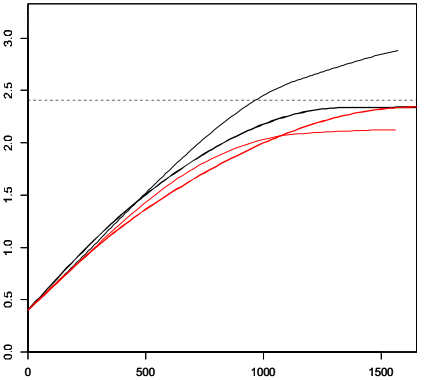


Illustration 2



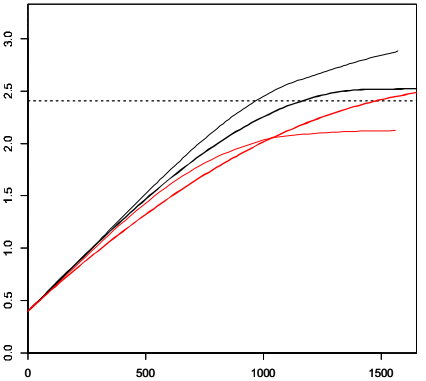
Geovariances

Illustration 3



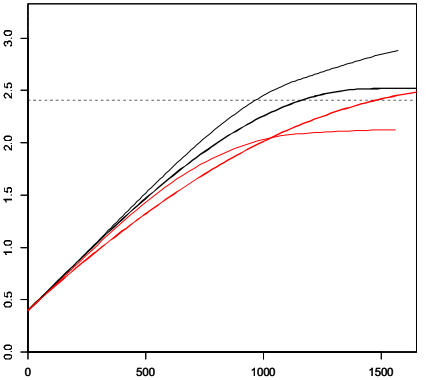
Geovariances

Illustration 4



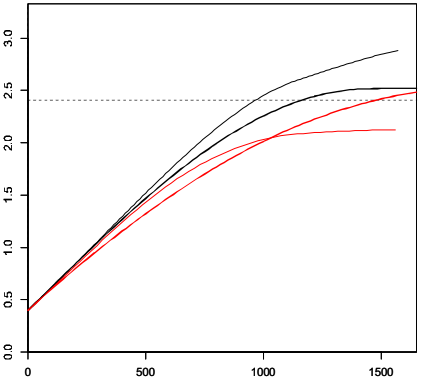
Geovariances

Illustration 5



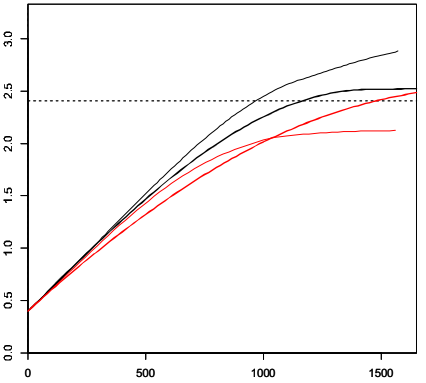
Geovariances

Illustration 6



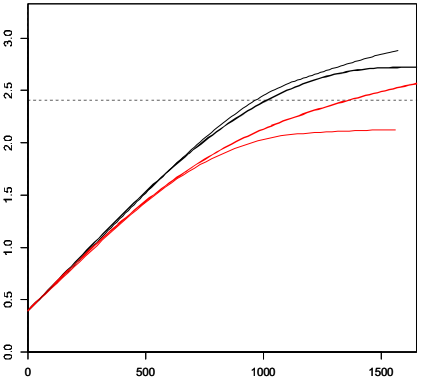
Geovariances

Illustration 7



Geovariances

Illustration 8



Geovariances

Illustration 9

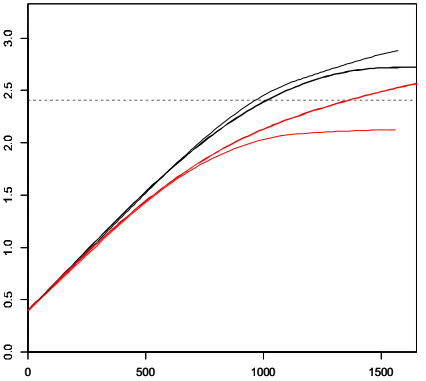
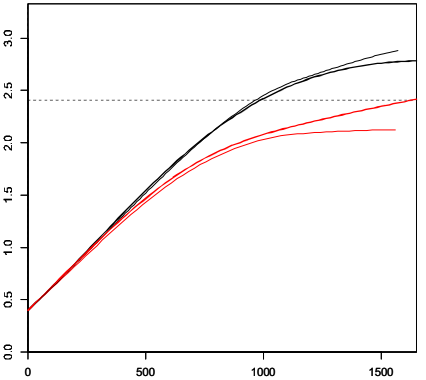


Illustration 10



Geovariances

Illustration 11

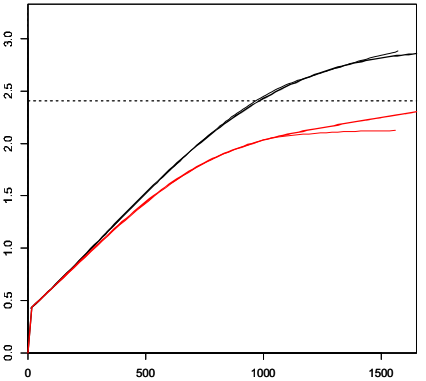


Illustration 12

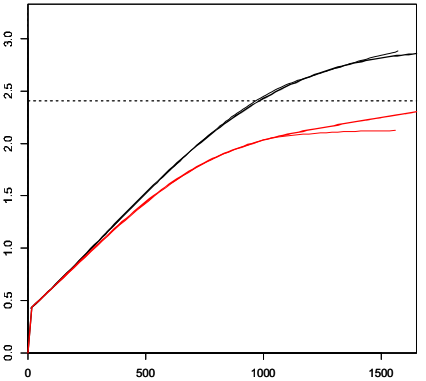
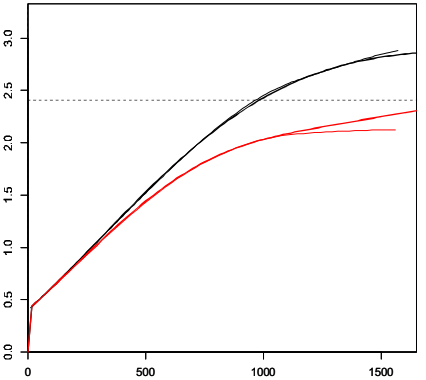


Illustration 13



Geovariances

Illustration 14

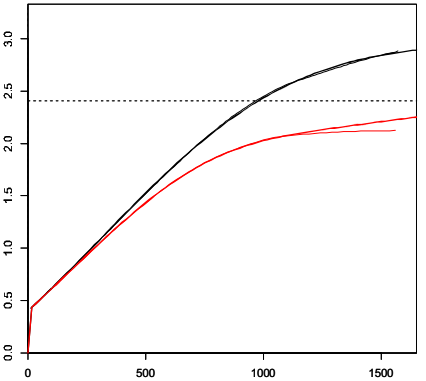
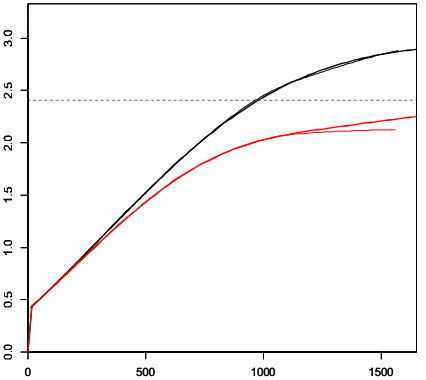
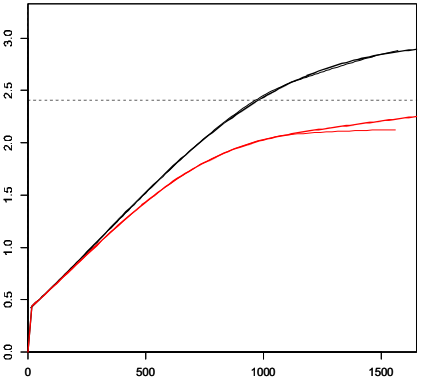


Illustration 15



Geovariances

Illustration 16



Geovariances

Illustration 17

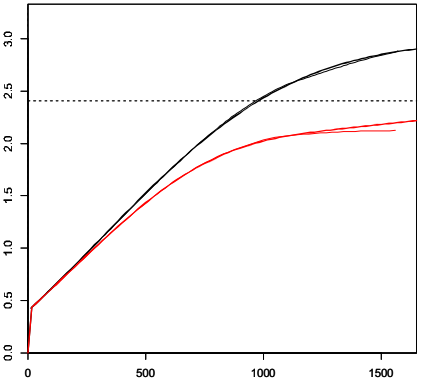
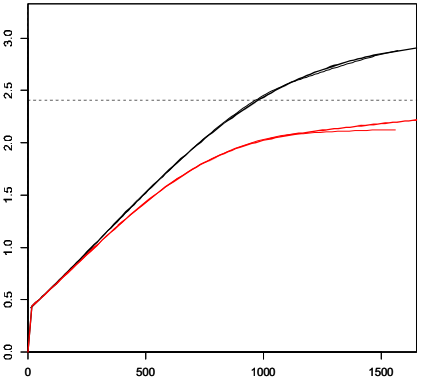


Illustration 18



Geovariances

Illustration 19

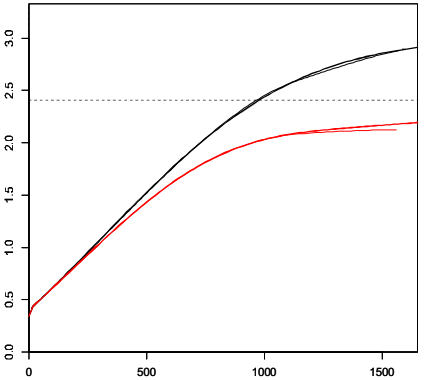
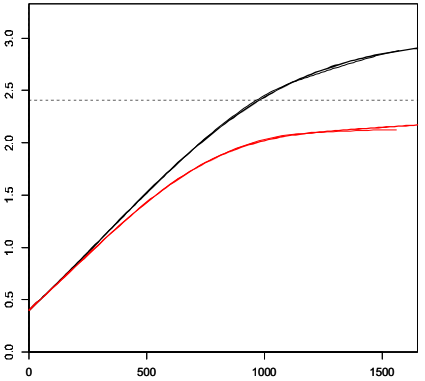
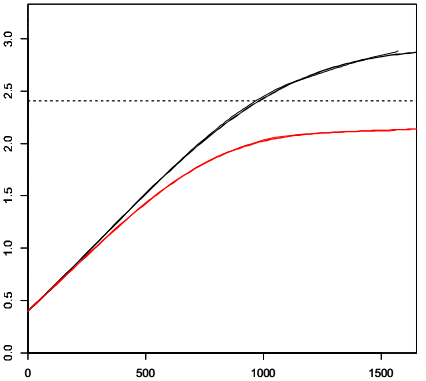


Illustration 20



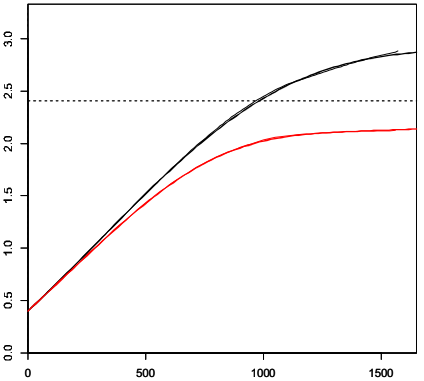
Geovariances

Illustration 21



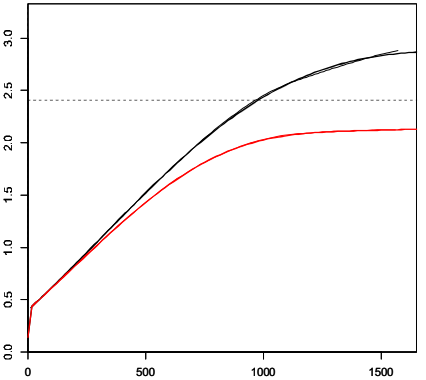
Geovariances

Illustration 22



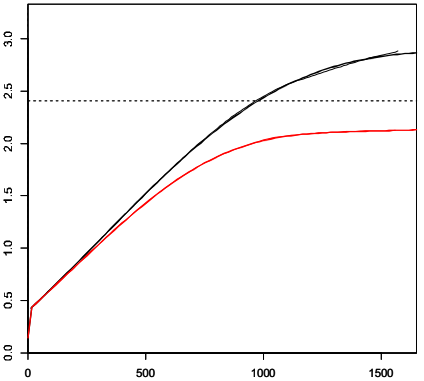
Geovariances

Illustration 23

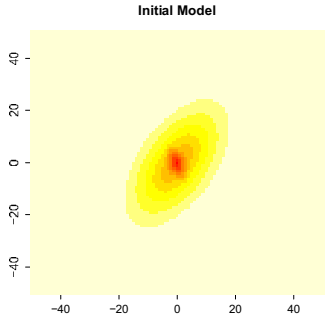


Geovariances

Illustration 24

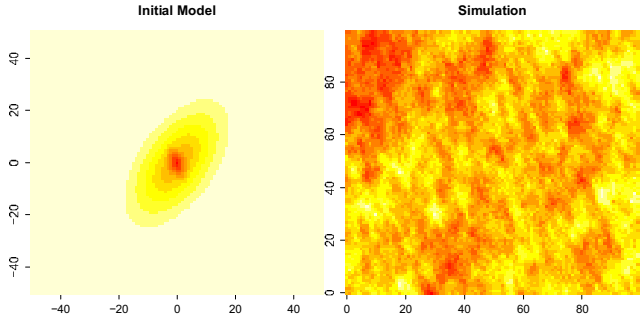


Variogram map



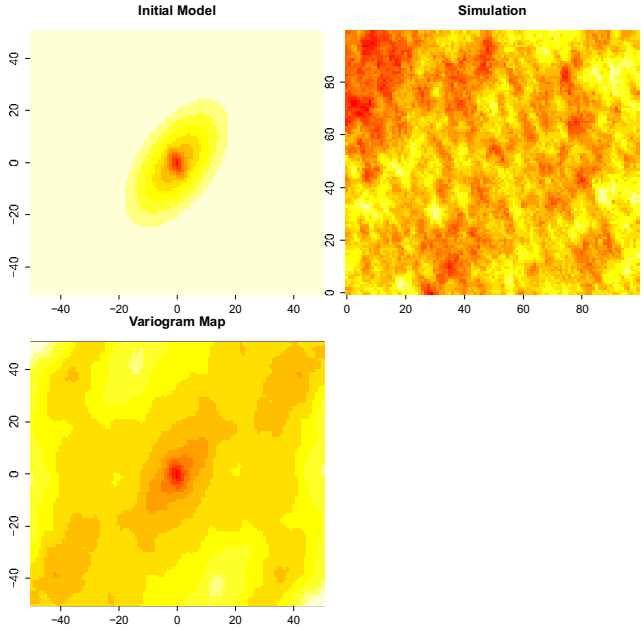
Geovariances

Variogram map



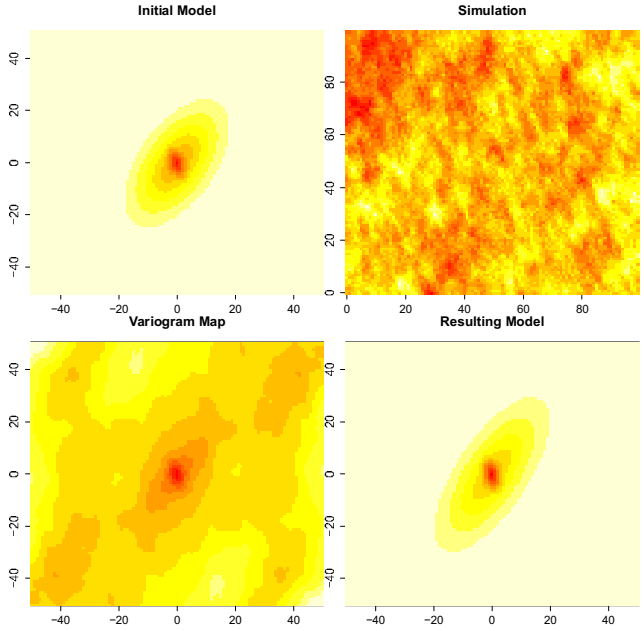
Geovariances

Variogram map



Gecvariances

Variogram map



Gecvariances

Linear model of coregionalisation

- K variables
- $\hat{\gamma}^{(ik)}(\vec{h}_j), i, k = 1, \ldots, K$ cross-variograms
- p structures $\gamma_{\theta_r}, r = 1, \ldots, p$
- Model :

$$\gamma^{(ik)}(\vec{h}) = \sum_{r=1}^{p} \lambda_{ik}^{(r)} \gamma_{\theta_r}(\vec{h}),$$

- Minimize

$$S(\psi) = \sum_{j=1}^{n} \sum_{i,k=1}^{K} \omega_{ikj} \left(\gamma_{\psi}^{(ik)}(\vec{h}_j) - \hat{\gamma}^{(ik)}(\vec{h}_j) \right)^2$$

with each $\Gamma^{(r)} = (\lambda_{ik}^{(r)})_{ik}$ for $i, k = 1, \ldots, K$ positive-definite

Profiled cost-function

- $\psi = (\Theta, \Lambda)$ and $S(\psi) = S(\Theta, \Lambda)$

Profiled cost-function

- $\psi = (\Theta, \Lambda)$ and $S(\psi) = S(\Theta, \Lambda)$
- For a given Θ sills fitting with the algorithm of Goulard and Voltz (1992)

$$\Lambda^\star(\Theta) = \underset{\Lambda \in \mathcal{D}_\Lambda}{\text{argmin}}\, S(\Theta, \Lambda)$$

Profiled cost-function

- $\psi = (\Theta, \Lambda)$ and $S(\psi) = S(\Theta, \Lambda)$
- For a given Θ sills fitting with the algorithm of Goulard and Voltz (1992)

$$\Lambda^\star(\Theta) = \underset{\Lambda \in \mathcal{D}_\Lambda}{\text{argmin}}\, S(\Theta, \Lambda)$$

- Profiled cost-function (reduction of the dimension)

$$\begin{aligned} S_r(\Theta) &= S(\Theta, \Lambda^\star(\Theta)) \\ &= \sum_{j=1}^{n} \sum_{i,k=1}^{K} \omega_{ijk} \left(\gamma^{(ik)}_{(\Theta, \Lambda^\star(\Theta))}(\vec{h_j}) - \hat{\gamma}^{(ik)}(\vec{h_j}) \right)^2 \end{aligned}$$

Profiled cost-function

- $\psi = (\Theta, \Lambda)$ and $S(\psi) = S(\Theta, \Lambda)$
- For a given Θ sills fitting with the algorithm of Goulard and Voltz (1992)

$$\Lambda^\star(\Theta) = \underset{\Lambda \in \mathcal{D}_\Lambda}{\text{argmin}}\, S(\Theta, \Lambda)$$

- Profiled cost-function (reduction of the dimension)

$$\begin{aligned} S_r(\Theta) &= S(\Theta, \Lambda^\star(\Theta)) \\ &= \sum_{j=1}^{n} \sum_{i,k=1}^{K} \omega_{ijk} \left(\gamma^{(ik)}_{(\Theta, \Lambda^\star(\Theta))}(\vec{h_j}) - \hat{\gamma}^{(ik)}(\vec{h_j}) \right)^2 \end{aligned}$$

- Foxleg gives

$$\Theta^\star = \underset{\Theta \in \mathcal{D}_\Theta}{\text{argmin}}\, S_r(\Theta)$$

Profiled cost-function

- $\psi = (\Theta, \Lambda)$ and $S(\psi) = S(\Theta, \Lambda)$
- For a given Θ sills fitting with the algorithm of Goulard and Voltz (1992)

$$\Lambda^\star(\Theta) = \underset{\Lambda \in \mathcal{D}_\Lambda}{\text{argmin}}\, S(\Theta, \Lambda)$$

- Profiled cost-function (reduction of the dimension)

$$\begin{aligned} S_r(\Theta) &= S(\Theta, \Lambda^\star(\Theta)) \\ &= \sum_{j=1}^{n} \sum_{i,k=1}^{K} \omega_{ijk} \left(\gamma^{(ik)}_{(\Theta, \Lambda^\star(\Theta))}(\vec{h_j}) - \hat{\gamma}^{(ik)}(\vec{h_j}) \right)^2 \end{aligned}$$

- Foxleg gives

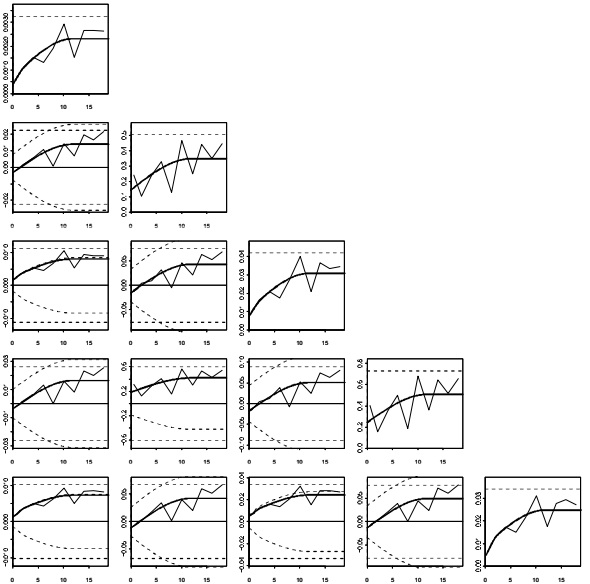
$$\Theta^\star = \underset{\Theta \in \mathcal{D}_\Theta}{\text{argmin}}\, S_r(\Theta)$$

- We get

$$\psi^\star = (\Theta^\star, \Lambda^\star(\Theta^\star)).$$

Gecvariances

18 variables and 5 structures



Gecvariances

Thank for your attention.

Desassis and Renard (2012) **Mathematical Geosciences**
Automatic Variogram Modeling by Iterative Least Squares : Univariate and Multivariate Cases

Geovariances