Kriging for processes solving partial differential equations

Karl Gerald van den Boogaart*

1st July 2001

Abstract

Physical laws are often expressed in terms of partial differential equations. However these laws completely determine the observable process only, when sufficiently many boundary conditions are known. We can use numerical methods like splines or finite elements to interpolate such processes.

Kriging is a method to interpolate stationary random processes based on their estimated second order moments, typically unaware of the physical law governing the observed spatial process.

For linear partial differential equations both approaches can be unified. It can be shown that linear differential equations impose restrictions on the class of admissible variograms and trend models. Thus the known physical law can help to select a physically reasonable variogram model. When these restrictions are honored, then the resulting universal kriging estimates and conditional simulations solve the differential equation in mean square sense.

As a introductory example the problem of kriging the gravity potential of the earth in free space is considered and it is shown that none of the commonly used variogram models is admissible. General methods to construct physically admissible variograms are shown.

1 Introduction

1.1 Motivation

The processes and structures of a landscape and an ore body as investigated by geology are often well understood in a qualitative manner. Most

*Mathematics and Computer Sciences in Geology, Freiberg University of Mining and Technology, Germany

phenomena can be related to physical laws. However, our incomplete knowledge about the relevant physical processes, unknown boundary conditions and superposition of the simple phenomenon by noisy irregularities makes it impossible to understand the landscape as the unique solution of a single differential equation given by the one relevant physical law. Thus we often have to restrict our modeling of landscape to simple and well investigated areas or we need to utilize purely statistical techniques like kriging. In kriging the landscape is understood as the result of a random genesis of the land described only by the phenomenological covariance function. From the theory of stochastic differential equations (e.g. [Gard 1988]) we know that a stochastic differential equation imposes some restrictions on the covariogram function of the process. Especially the Laplace differential equation has often been discussed in the context of kriging (e.g. [Chilès&Delfiner 1999, Christakos 1992]).

This contribution summarizes some simple and useful results concerning kriging of processes solving linear partial differential equations. Although most of them might be obvious and surely well known in some branches of science, I found it difficult to find explicite statements in the context of kriging of the simple facts, which this contribution is based on. Thus no citations does not necessaryly mean that there is none, but that I found non.

1.2 Notations of Kriging

Shortly speaking: Kriging solves the problem to estimate the local value $f(\mathbf{x})$ at the location $\mathbf{x}$ of a location dependent quantity from measurements $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)$ of this quantity at some other locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$ as good as possible by a linear estimation function:

$$\text{Estimation for } f(\mathbf{x}) = \sum_{i=1}^{n} \lambda_i f(\mathbf{x}_1) \tag{1}$$

Typical quantities are density, temperature, concentration, potential, elevation over sea level or local thickness of a deposit. The local values of the quantity are modeled as a random function. The location can be a three dimensional position or a geographic location.

We need to know a linear model for the expected value $E[f(\mathbf{x})]$ of the quantity

$$E[f(\mathbf{x})] = g_0(\mathbf{x}) + \sum_{k=1}^{m} \beta_k g_k(\mathbf{x}) \tag{2}$$

Here the $g_k(\mathbf{x}), k = 0, \ldots, m$ are some known functions and the $\beta_k, k = 1, \ldots, m$ are unknown real valued parameters. Further we need to know the

covariance of the values of the quantity at different locations:

$$c(\mathbf{x}, \mathbf{y}) = \mathrm{cov}(f(\mathbf{x}), f(\mathbf{y})) \tag{3}$$

The function $c(\mathbf{x}, \mathbf{y})$ is called a covariogram. The kriging estimation for $f(\mathbf{x})$ is than given by formula (11). A standard estimation error is provided for this estimate. For some trend models, where the $g_1(\mathbf{x})$ is a constant the covariogram can be replaced by the variogram $\gamma(\mathbf{x}, \mathbf{y}) := \mathrm{var}(f(\mathbf{x}) - f(\mathbf{y}))$. The covariogram and the variogram are typically estimated from the observed values $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n)$. Typically it is assumed that $c(\mathbf{x}, \mathbf{y})$ only depends on the vectorial distance $\mathbf{x} - \mathbf{y}$ of the locations $\mathbf{x}$ and $\mathbf{y}$: $c(\mathbf{x}, \mathbf{y}) = c_s(\mathbf{x} - \mathbf{y})$. This assumption is called (second order) stationarity. For a more detailed introduction to kriging the reader is referred to [Wackernagel 1998, Chilès&Delfiner 1999, Cressie 1993].

1.3 Differential Equations in Mean Square Sense

1.3.1 Stochastic Differentiation

To consider differential equations of random fields we need the notion of differentiation in mean square sense (e.g. [Christakos 1992]): A random function $f : \mathbb{R}^d \to \mathbb{R}$ is called mean square differentiable in $\mathbf{x}_0$ in direction $\mathbf{x}$ when the differential quotient

$$\lim_{h \to 0} \frac{f(\mathbf{x}_0 + h\mathbf{x}) - f(\mathbf{x}_0)}{h}$$

has a limit of the mean and the variance for $h \to 0$ [Chilès&Delfiner 1999]:

$$\left| \lim_{h \to 0} E\left(\frac{f(\mathbf{x}_0 + h\mathbf{x}) - f(\mathbf{x}_0)}{h} \right) \right| < \infty \text{ and } \lim_{h \to 0} \mathrm{var}\left(\frac{f(\mathbf{x}_0 + h\mathbf{x}) - f(\mathbf{x}_0)}{h} \right) < \infty$$

High order derivatives, partial derivatives and gradients are defined analogously to derivatives of deterministic functions. When

$$\nabla_{\mathbf{x}} f(\mathbf{x}) := \left(\ \frac{\partial}{\partial x_1} f(\mathbf{x}), \ldots, \frac{\partial}{\partial x_d} f(\mathbf{x}) \ \right)^t$$

exists and has finite variation it can be shown, that[Chilès&Delfiner 1999]

$$E\left(\nabla_{\mathbf{x}} f(\mathbf{x})\right) = \nabla_{\mathbf{x}} E[f(\mathbf{x})]$$

$$\mathrm{cov}\left(\nabla_{\mathbf{x}} f(\mathbf{x}), f(\mathbf{y})\right)) = \nabla_{\mathbf{x}} c(\mathbf{x}, \mathbf{y}) \tag{4}$$

$$\mathrm{cov}\left(\nabla_{\mathbf{x}} f(\mathbf{x}), \nabla_{\mathbf{y}} f(\mathbf{y})\right)) = \nabla_{\mathbf{x}} \nabla_{\mathbf{y}} c(\mathbf{x}, \mathbf{y}) \tag{5}$$

Where $c(\mathbf{x},\mathbf{y}) = \mathrm{cov}(f(\mathbf{x}), f(\mathbf{y}))$ denotes the covariogram of the process $f(\mathbf{x})$. Note that a function is by definition n times differentiable in mean square sense, if and only if the derivative $\frac{\mathrm{d}^n}{\mathrm{d}\mathbf{x}^n}\frac{\mathrm{d}^n}{\mathrm{d}\mathbf{x}^n}c(\mathbf{x},\mathbf{y})|_{x=y}$ of the covariogram exists[Christakos 1992]. From relations seen later this corresponds to the existence of $\frac{\mathrm{d}^{2n}}{\mathrm{d}\mathbf{h}^{2n}}c(\mathbf{h})|_{\mathbf{h}=0}$ in case of isotropy. This relation is often stated in literature especially for the first derivative (e.g. [Chilès&Delfiner 1999]).

1.3.2 The General Linear Partial Differential Equation

In this paper we only consider linear partial differential equations of the type:

$$L_{\mathbf{x}}f(\mathbf{x}) := \sum_{i=0}^{n} \alpha_{ij}^{j_1\ldots j_i}(\mathbf{x}) \frac{\partial^i}{\partial x_{j_1}\ldots\partial x_{j_i}} f(\mathbf{x}) = \mathbf{k}_j(\mathbf{x}) \tag{6}$$

with some tensor valued functions $\alpha_i : \mathbb{R}^d \to \mathbb{R}^{p \times \overbrace{d\times\ldots\times d}^{i\times}}$, $\mathbf{k} : \mathbb{R}^d \to \mathbb{R}^p$, $f : \mathbb{R}^d \to \mathbb{R}$. Summation over equal indices $j_1, \ldots, j_n$ is implied according to the sum convention for tensors. The location $\mathbf{x} \in \mathbb{R}^d$ can be a location in space or a location in space and time.

We call an equation homogeneous, if $\mathbf{k}(\mathbf{x}) = 0$. We call it stationary when the α_i and $\mathbf{k}$ do not depend on $\mathbf{x}$. We call it isotrope when the equation is invariant under orthogonal transformations of the coordinate system (which implies stationarity and eveness). We call it even if $L_{\mathbf{x}}f(\mathbf{x}) = L_{\mathbf{x}}f(-\mathbf{x})$ for all even functions f. This is equivalent to $\alpha_i \equiv 0$ for i odd.

The concepts on stationarity and isotropy are closely related to these concepts in geostatistics. As an example we consider the Laplace equation:

$$\Delta_{\mathbf{x}}f(\mathbf{x}) = (\frac{\mathrm{d}^2}{\mathrm{d}x_1^2} + \frac{\mathrm{d}^2}{\mathrm{d}x_2^2}) = 0$$

in two dimensions, which is homogeneous, stationary, even and isotrope.

We start up from a well known theorem (e.g. [Christakos 1992]):

Theorem 1 (Moments of linear transformed processes) *For two random functions $f(\mathbf{x})$ and $g(\mathbf{y})$ for which $E[f(\mathbf{x})]$, $E[g(\mathbf{x})]$, $\mathrm{cov}(f(\mathbf{x}), g(\mathbf{y}))$ exist and are all finite, and for any differential operator $L_{\mathbf{x}}$ such that $L_{\mathbf{x}}E[f(\mathbf{x})]$ and $L_{\mathbf{x}}\mathrm{cov}(f(\mathbf{x}), g(\mathbf{y}))$ exist and are finite it holds in mean square sense:*

$$E[L_{\mathbf{x}}f(\mathbf{x})] = L_{\mathbf{x}}E[f(\mathbf{x})] \tag{7}$$

$$\mathrm{cov}(L_{\mathbf{x}}f(\mathbf{x}), g(\mathbf{y})) = L_{\mathbf{x}}\mathrm{cov}(f(\mathbf{x}), g(\mathbf{y})) \tag{8}$$

2 Restrictions on the Covariogram

From the theory of stochastic differential equation (e.g. [Gard 1988]) we know that a linear equation imposes some restrictions on the moments of the distribution. The implications of these restrictions with respect to the different typical assumption used with kriging are summarized here.

2.1 Instationary Covariograms

Suppose our process solves $L_{\mathbf{x}} f(\mathbf{x}) = \mathbf{k}(\mathbf{x})$. Since $\mathbf{k}(\mathbf{x})$ does not depend on the realization of the process f also $L_{\mathbf{x}} f(\mathbf{x})$ does not depend on it and thus has variance 0.

$$0 = \mathrm{var}(L_{\mathbf{x}} f(\mathbf{x})) = L_{\mathbf{x}} L_{\mathbf{y}} c(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\mathbf{y}}$$

and thus

$$0 = \mathrm{cov}(L_{\mathbf{x}} f(\mathbf{x}), f(\mathbf{y})) = L_{\mathbf{x}} c(\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{y}$$

since the covariance with something must be zero when the variance of it is zero. Further we know $L_{\mathbf{x}} f(\mathbf{x}) = \mathbf{k}(\mathbf{x})$ and thus

$$L_{\mathbf{x}} E[f(\mathbf{x})] = E[L_{\mathbf{x}} f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$$

and on the other hand with $f_v(\mathbf{x}) := (f(\mathbf{x}) - E[f(\mathbf{x})])$

$$L_{\mathbf{x}} f_v(\mathbf{x}) = 0$$

Since with $E[f_v(\mathbf{x})] = 0 \forall x$ all its derivatives have zero expectation, too, we get

$$L_{\mathbf{x}} E[f_v(\mathbf{x})] = 0$$

Moreover since

$$\mathrm{var}(L_{\mathbf{x}} f_v(\mathbf{x})) = L_{\mathbf{x}} L_{\mathbf{y}} c(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$$

we have

$$L_{\mathbf{x}} f_v(\mathbf{x}) = E[L_{\mathbf{x}} f_v(\mathbf{x})] = L_{\mathbf{x}} E[f_v(\mathbf{x})] = 0$$

This all leads to the useful corollary, which will be extended to more complex situations later:

Theorem 2 *The following three conditions are equivalent:*

1. $f(\mathbf{x})$ *solves* $L_{\mathbf{x}} f(\mathbf{x}) = \mathbf{k}(\mathbf{x})$ *in mean square sense.*
2. *The following two conditions hold simultaneously:*
   - $L_{\mathbf{x}} E[f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$

- $L_{\mathbf{x}}L_{\mathbf{y}}c(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$

3. *The following two conditions hold simultaneously:*

   - $L_{\mathbf{x}}E[f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$
   - $L_{\mathbf{x}}c(\mathbf{x},\mathbf{y}) = 0 \forall \mathbf{y}$

Remark 1 *Simular relations hold for the variogram of* $f(\mathbf{x})$*:*

$$L_{\mathbf{x}}\gamma(\mathbf{x},\mathbf{y}) = L_{\mathbf{x}}\left(c(\mathbf{x},\mathbf{x}) + c(\mathbf{y},\mathbf{y}) - 2c(\mathbf{x},\mathbf{y})\right) = 0 \forall \mathbf{y}$$

$$L_{\mathbf{x}}L_{\mathbf{y}}\gamma(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$$

2.2 Stationarity

It does generally not make any sense to have a stationary variogram or covariogram and a non stationary differential equation since then we have a stationary covariogram function solving infinitely many differential equations. Thus let us consider the case of stationary variogram and stationary differential equation. We want to rewrite the variogram and covariogram as:

$$\gamma(\mathbf{x},\mathbf{y}) := \gamma_s(\mathbf{x}-\mathbf{y}), \quad c(\mathbf{x},\mathbf{y}) := c_s(\mathbf{x}-\mathbf{y})$$

We get

$$L_{\mathbf{x}}L_{\mathbf{y}}\gamma(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = L_{\mathbf{x}}L_{\mathbf{y}}\gamma_s(\mathbf{y}-\mathbf{x})|_{\mathbf{x}=\mathbf{y}} = L_{\mathbf{x}}(L_{\mathbf{h}}\gamma_s)(\mathbf{y}-\mathbf{x})|_{\mathbf{x}=\mathbf{y}}$$

To get useful results we additionally have to assume that $L_{\mathbf{x}}$ is even since then we obtain:

$$L_{\mathbf{x}}(L_{\mathbf{h}}\gamma_s)(\mathbf{y}-\mathbf{x})|_{\mathbf{x}=\mathbf{y}} = L_{\mathbf{x}}(L_{\mathbf{h}}\gamma_s)(\mathbf{x}-\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = (L_{\mathbf{h}}L_{\mathbf{h}}\gamma_s)(0)$$

and in the same way we get

$$L_{\mathbf{x}}L_{\mathbf{y}}c(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = (L_{\mathbf{h}}L_{\mathbf{h}}c_s)(0)$$

Thus we can specialize corollary 2 to

Theorem 3 *If* $L_{\mathbf{x}}$ *is stationary and even the following two conditions are equivalent with covariogram* $c(\mathbf{x},\mathbf{y}) = c_s(\mathbf{x}-\mathbf{y})$*:*

1. $f(\mathbf{x})$ *solves* $L_{\mathbf{x}}f(\mathbf{x}) = \mathbf{k}(\mathbf{x})$ *in mean square sense.*

2. *The following two conditions hold simultaneously:*

- $L_{\mathbf{x}}E[f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$
- $(L_{\mathbf{h}}L_{\mathbf{h}}c_s)(0) = 0$

3. The following two conditions hold simultaneously:

- $L_{\mathbf{x}}E[f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$
- $(L_{\mathbf{h}}c_s)(0) = 0 \,\forall \mathbf{h}$

Thus the differential equation imposes a finite number of conditions on the derivatives of c_s in the origin.

2.2.1 Restrictions on the Stationary Variogram

The variogram is just a affine linear transformation of the covariogram

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

All its derivatives are the negative of the corresponding derivatives of c. Thus when the differential equation does not contain terms containing the zero order derivative, that is it does not depend on $f(\mathbf{x})$ itself, the same restrictions apply to the variogram as to the covariogram, since the homogeneous linear differential equation is invariant with respect to the sign of the derivatives. For details see section 6.

2.3 Isotropy

Imposing the same arguments as for stationarity we should only consider isotropy of the differential equation and isotropy of the variogram (covariogram) at the same time. We want to rewrite $c(\mathbf{x}, \mathbf{y})$ as

$$c(\mathbf{x}, \mathbf{y}) = c_I(\|\mathbf{x} - \mathbf{y}\|)$$

In order to get simple relations on $c_I(h),\ h \in \mathbb{R}$ we need a well known representation theorem for stationary and isotropic differential equations:

Theorem 4 (Δ-expansion of isotropic linear pde) *If*

$$L_{\mathbf{x}}f(\mathbf{x}) := \sum_{i=0}^{n} \alpha_{ij}^{j_1 \dots j_i}(\mathbf{x}) \frac{\partial^i}{\partial x_{j_1} \dots \partial x_{j_i}} f(\mathbf{x}) = \mathbf{k}_j(\mathbf{x})$$

with some tensor valued coefficients $\alpha^{(i)}(x) \in \mathbb{R}^{p \times \overbrace{d \times \dots \times d}^{i\times}}$. *is a stationary, homogeneous, isotropic, linear partial differential equation of order n,*

then it can be rewritten as:

$$\sum_{i=0}^{n/2} \beta^{(2i)} \Delta^i f(\mathbf{x}) = \mathbf{k}(\mathbf{x}),\ \beta_k^{(i)} := \sum_{j_1,\ldots,j_n=1}^{p} \alpha_{kj_1\ldots j_n}^{(i)}$$

Where $\Delta := \sum_{i=1}^{d} \frac{d^2}{dx_i^2}$ *is the Laplace operator.*

Without proof. Compare [Müller 1998].

To rewrite the differential equation $L_{\mathbf{x}} L_{\mathbf{y}} c(\|\mathbf{x}-\mathbf{y}\|)|_{x=y} = 0$ in terms of β we first consider

$$\begin{aligned}
\Delta_{\mathbf{h}} c_I(\|\mathbf{h}\|) &= \sum_{i=1}^{d} \frac{d^2}{dx_i^2} c_I\left(\|\mathbf{h}\|\right) = \\
&= \sum_{i=1}^{d} \frac{d}{dx_i} \left(c_I'\left(\|\mathbf{h}\|\right) \frac{x_i}{\|\mathbf{h}\|} \right) = \\
&= \sum_{i=1}^{d} \left(c_I''\left(\|\mathbf{h}\|\right) \frac{x_i^2}{\|\mathbf{h}\|^2} + c_I'\left(\|\mathbf{h}\|\right) \frac{\|\mathbf{h}\| - \frac{x_i^2}{\|\mathbf{h}\|}}{\|\mathbf{h}\|^2} \right) = \\
&= c_I''\left(\|\mathbf{h}\|\right) + (d-1) \frac{c_I'\left(\|\mathbf{h}\|\right)}{\|\mathbf{h}\|} \qquad (9)
\end{aligned}$$

And thus according to the rule of L'Hospital since $c_I'(0) = 0$ due to isotropy we get:

$$\Delta_{\mathbf{h}} c_I(\|\mathbf{h}\|)|_{\mathbf{h}=0} = d c_I''(0)$$

From recursive application of formula (9) on the Taylor series of $c_I(h)$ we get explicitly for high order monomials in Δ:

$$\Delta_{\mathbf{h}}^i c_I(\|\mathbf{h}\|)|_{\mathbf{h}=0} = \left(\frac{d^{2i}}{dh^{2i}} c_I(h)|_{h=0} \right) \prod_{j=1}^{i} \left(\frac{(2j+d-2)}{(2j-1)} \right) \qquad (10)$$

Proof: Note that the Taylor series

$$c_I(\|\mathbf{h}\|) = \sum_{i=0}^{2n} a_i \|\mathbf{h}\|^i + o(|\mathbf{h}\|^{2n+1})$$

of $c(h)$ must have only even Taylor coefficients up to order $2n$ since otherwise it could not solve the differential equation since $c_I(\|\mathbf{h}\|)$ would not be $2n$ times differentiable in $h = 0$ and thus $c(\|\mathbf{h}\|)$ could not be $2n$ times differentiable in $\mathbf{h} = 0$ and thus $f(x)$ could not be n times differentiable in mean square sense.

Applying one step of equation (9) would transform this to

$$\Delta c_I(\|h\|) = \sum_{i=2}^{\infty} \left(\underbrace{a_i i(i-1)}_{c_I''} + \underbrace{a_i i(d-1)}_{\frac{c_I'}{\|h\|}} \right) \|h\|^{i-2}$$

Index transformation, recursive application and comparison to the Taylor series of $\frac{d^{2n}}{dh^{2n}} c(h)$ yields formula (10).

Thus we can specialize corollary 2 further:

Theorem 5 *Let $c(\mathbf{x}, \mathbf{y}) = c(\|\mathbf{x} - \mathbf{y}\|)$ be the covariogram of an isotropic random process $f(\mathbf{x})$. Then $f(\mathbf{x})$ solves the equation*

$$\sum_{i=0}^{n/2} \beta^{(2i)} \Delta^i f(\mathbf{x}) = \mathbf{k}(\mathbf{x})$$

in mean square sense if and only if

$$\sum_{i=0}^{n/2} \beta^{(2i)} \Delta^i E[f(\mathbf{x})] = \mathbf{k}(\mathbf{x})$$

and

$$\sum_{i=0}^{n} \eta^{(2i)} \frac{d^{2i}}{dh^{2i}} c(h)|_{h=0} = 0$$

with

$$\eta_p^{2i} = \left(\prod_{j=1}^{i} \frac{(2j+d-2)}{(2j-1)} \right) \sum_{j=0}^{i} \beta_p^{(2j)} \beta_p^{(2i-2j)}$$

Summarizingly, any linear isotropic differential equation of degree n induces linear constraints on the derivatives of $c(h)$ at $h = 0$ up to degree $2n$.

Thus we can detect the qualitative behavior of the variogram near the origin from the differential equation. This is particularly useful, because the kriging weights often depend only on the behavior of the variogram near the origin.

2.4 Example

Let us consider the example of the Laplace differential equation $\Delta_{\mathbf{x}} f(\mathbf{x}) = 0$ and an isotropic variogram model $\gamma(h)$. The conditions are:

$$\begin{aligned} 0 &= \Delta_{\mathbf{h}} \Delta_{\mathbf{h}} \gamma(\|\mathbf{h}\|) = \\ &= \Delta_{\mathbf{h}}^2 c(\sqrt{x_1^2 + x_2^2}) \\ &= \frac{d^4}{dx_1^4} + \frac{d^4}{dx_2^4} + 2 \frac{d^2}{dx_1^2} \frac{d^2}{dx_1^2} c(\sqrt{x_1^2 + x_2^2}) \end{aligned}$$

Using corollary 5 we can rewrite it:

$$0 = \left(\prod_{j=1}^{2} \frac{(2j+d-2)}{(2j-1)} \right) \gamma''''(h) = \frac{d(d+2)}{2} \gamma''''(h)$$

Thus we just need the fourth derivate of γ in the origin to be zero. This condition is not met by any of the commonly used Gaussian, exponential or spherical variograms. Thus we later have to find out how to construct variograms and covariograms with such properties.

3 Kriging Estimators Solve the Differential Equation

The implications on the covariogram are interesting for kriging in two ways:

- They help to find the correct kriging weights, since they help to find an appropriate variogram model.
- They imply that the kriging results solve the differential equations.

3.1 The Kriging Interpolation Solves the Differential Equation

Theorem 6 *The universal kriging interpolation*

$$\hat{f}(x) = g_0(\mathbf{x}) + \begin{pmatrix} f(\mathbf{x}_1) - g_0(\mathbf{x}_1) \\ \vdots \\ f(\mathbf{x}_m) - g_0(\mathbf{x}_m) \\ 0 \\ \vdots \\ 0 \end{pmatrix}^t \cdot \begin{pmatrix} c(\mathbf{x}_1,\mathbf{x}_1) & \cdots & c(\mathbf{x}_1,\mathbf{x}_m) & g_1(x_1) & \cdots & g_p(x_1) \\ \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ c(\mathbf{x}_m,\mathbf{x}_1) & \cdots & c(\mathbf{x}_m,\mathbf{x}_m) & g_1(x_m) & \cdots & g_p(x_m) \\ g_1(x_1) & \cdots & g_1(x_m) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_p(x_1) & \cdots & g_p(x_m) & 0 & \cdots & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} c(\mathbf{x},\mathbf{x}_1) \\ \vdots \\ c(\mathbf{x},\mathbf{x}_m) \\ g_1(\mathbf{x}) \\ \vdots \\ g_p(\mathbf{x}) \end{pmatrix} \tag{11}$$

solves the linear partial differential equation

$$L_{\mathbf{x}} f(\mathbf{x}) = k(\mathbf{x})$$

if the covariogram is admissible for that differential equation

$$L_{\mathbf{x}} L_{\mathbf{y}} c(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0,$$

the trend function $g_1, \ldots, g_p$ solve the homogeneous equation

$$L_{\mathbf{x}} g_i(\mathbf{x}) = 0 \forall i = 1, \ldots, p$$

and the fixed trend part g_0 solves the heterogeneous differential equation

$$L_{\mathbf{x}} g_0(\mathbf{x}) = k(\mathbf{x}).$$

Proof: For any fixed realization $f(\mathbf{x}_1), \ldots, f(\mathbf{x_m})$ we can write

$$\hat{f}(\mathbf{x}) = g_0(\mathbf{x}) + \sum_{i=1}^{m} q_i c(\mathbf{x}, \mathbf{x}_m) + \sum_{i=1}^{p} r_i g(\mathbf{x})$$

since the matrix and the row vector in the formula for $\hat{f}(\mathbf{x})$ do not depend on $\mathbf{x}$ but only on the $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and the realization. We get

$$L_{\mathbf{x}} \hat{f}(x) = \underbrace{L_{\mathbf{x}} g_0(\mathbf{x})}_{k(\mathbf{x})} + \sum_{i=1}^{m} q_i \underbrace{L_{\mathbf{x}} c(\mathbf{x}, \mathbf{x}_m)}_{0} + \sum_{i=1}^{p} r_i \underbrace{L_{\mathbf{x}} g(\mathbf{x})}_{0} = k(\mathbf{x})$$

The $L_{\mathbf{x}} c(\mathbf{x}, \mathbf{x}_m) = 0$ follows from Theorem 2.

In general $c(\mathbf{x}, \mathbf{y})$ can be replaced by any generalized covariogram with respect to the trend functions $g(\mathbf{x})$ without changing the results. Especially, if $g_1(\mathbf{x}) \equiv 1$ solves the homogeneous differential equation the variogram can be substituted for the covariogram, see the proof in section 6.

3.2 Simulated Realizations

A second important application of kriging is that we can simulate realizations of the processes conditional on the observed quantities. This is important in applications, since it helps to understand the variability of possible realizations and to calculate probabilities for specific situations[Cressie 1993]. A simulation is a random function $\tilde{f}$. The covariance $\tilde{c}(\mathbf{x}, \mathbf{y})$ of this simulated random function depends on the observed locations. The conditional simulation based on universal kriging is given as:

$$\tilde{f}(\mathbf{x}) = \hat{f}(\mathbf{x}) + Z(\mathbf{x})$$

where $Z(\mathbf{x})$ is a random function with zero mean and covariogram $\tilde{c}(\mathbf{x}, \mathbf{y})$, which can be calculated from the covariogram. With abbreviations we get:

$$\mathbf{c}(x) := \begin{pmatrix} c(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ c(\mathbf{x}, \mathbf{x}_m) \end{pmatrix}$$

$$
\begin{aligned}
\mathbf{C} &:= \begin{pmatrix} c(\mathbf{x}_1,\mathbf{x}_1) & \cdots & c(\mathbf{x}_1,\mathbf{x}_m) \\ \vdots & \ddots & \vdots \\ c(\mathbf{x}_m,\mathbf{x}_1) & \cdots & c(\mathbf{x}_m,\mathbf{x}_m) \end{pmatrix} \\
\mathbf{g}(\mathbf{x}) &:= \begin{pmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_p(\mathbf{x}) \end{pmatrix} \\
\mathbf{G} &:= \begin{pmatrix} g_1(x_m) & \cdots & g_p(x_m) \\ \vdots & \ddots & \vdots \\ g_1(x_m) & \cdots & g_p(x_m) \end{pmatrix} \\
\mathbf{n}(\mathbf{x}) &:= \begin{pmatrix} \mathbf{c}(\mathbf{x}) \\ \mathbf{g}(\mathbf{x}) \end{pmatrix} \\
\mathbf{N} &:= \begin{pmatrix} \mathbf{C} & \mathbf{G} \\ \mathbf{G}^t & 0 \end{pmatrix} \\
\mathbf{T} &:= \left.\begin{pmatrix} 1 & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ & & 1 & 0 & \cdots & 0 \end{pmatrix}\right\} m \text{ columns} \\
\mathbf{w}(\mathbf{x}) &:= \mathbf{T}\mathbf{N}^{-1}\mathbf{n}(\mathbf{x})
\end{aligned}
$$

$$
\begin{aligned}
\tilde{c}(\mathbf{x},\mathbf{y}) &= \begin{pmatrix} 1 \\ 0 \\ -\mathbf{w}(\mathbf{x}) \end{pmatrix}^t \begin{pmatrix} c(\mathbf{x},\mathbf{x}) & c(\mathbf{x},\mathbf{y}) & \mathbf{c}(\mathbf{x})^t \\ c(\mathbf{y},\mathbf{x}) & c(\mathbf{y},\mathbf{y}) & \mathbf{c}(\mathbf{y})^t \\ \mathbf{c}(\mathbf{x}) & \mathbf{c}(\mathbf{y}) & \mathbf{C} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -\mathbf{w}(\mathbf{y}) \end{pmatrix} \\
&= c(\mathbf{x},\mathbf{y}) - \mathbf{w}(\mathbf{x})^t\mathbf{c}(\mathbf{y}) - \mathbf{c}(\mathbf{x})^t\mathbf{w}(\mathbf{y}) + \mathbf{w}(\mathbf{x})^t\mathbf{C}\mathbf{w}(\mathbf{y}) \\
&= c(\mathbf{x},\mathbf{y}) - \mathbf{n}^t(\mathbf{x})\mathbf{N}^{-1}\mathbf{T}^t\mathbf{c}(\mathbf{y}) - \mathbf{c}(\mathbf{x})^t\mathbf{T}\mathbf{N}^{-1}\mathbf{n}(\mathbf{y}) + \\
&\quad + \mathbf{n}^t(\mathbf{x})\mathbf{N}^{-1}\mathbf{T}^t\mathbf{C}\mathbf{T}\mathbf{N}^{-1}\mathbf{n}(\mathbf{y})
\end{aligned}
$$

Every term in this last formula is a linear superposition of functions solving

$$L_{\mathbf{x}} f(\mathbf{x}) = 0$$

and thus we have

$$L_{\mathbf{x}} \tilde{c}(\mathbf{x},\mathbf{y}) = 0 \; \forall \mathbf{y}$$

Together with $L_x \hat{f}(\mathbf{x}) = k(\mathbf{x})$ from theorem 6 we obtain from theorem 2 the following theorem:

Theorem 7 (Conditional simulations solves pde in mean square sense) *The conditional simulation*

$$\tilde{f}(\mathbf{x}) = \hat{f}(\mathbf{x}) + Z(\mathbf{x})$$

where $Z(\mathbf{x})$ is a random process with mean zero and covariogram

$$\tilde{c}(\mathbf{x},\mathbf{y}) = c(\mathbf{x},\mathbf{y}) - \mathbf{w}(\mathbf{x})^t\mathbf{c}(\mathbf{y}) - \mathbf{c}(\mathbf{x})^t\mathbf{w}(\mathbf{y}) + \mathbf{w}(\mathbf{x})^t\mathbf{C}\mathbf{w}(\mathbf{y})$$

solves

$$L_{\mathbf{x}}\tilde{f}(\mathbf{x}) = k(\mathbf{x})$$

in mean square sense.

4 Generalization to Tensors

For random fields, which are described by more than one number per location we need a multivariate extension of theorem 2. The full tensorial version of theorem 2 is given here. No extra proof is given, since it is just an extension of theorem 2 to tensorial notation. The tensorial notation implies summation over indices showing up twice in the same product (Einstein's summation convention).

Theorem 8 *For $\mathbf{f} : \mathbb{R}^d \to \mathbb{R}^{\times_{k=1}^N p_k}$ The following three conditions are equivalent:*

1. *$\mathbf{f}_{i_1\ldots i_n}(\mathbf{x})$ solves $\mathbf{L}^{i_1\ldots i_n}_{\mathbf{x}j_1\ldots j_M}\mathbf{f}_{i_1\ldots i_n}(\mathbf{x}) = \mathbf{k}_{j_1\ldots j_M}(\mathbf{x})$ in mean square sense.*
2. *The following two conditions hold simultaneously:*
   - $\mathbf{L}^{i_1\ldots i_N}_{\mathbf{x}j_1\ldots j_M}E[\mathbf{f}_{i_1\ldots i_N}(\mathbf{x})] = \mathbf{k}_{j_1\ldots j_M}(\mathbf{x})$
   - $\mathbf{L}^{i_1\ldots i_N}_{\mathbf{x}j_1\ldots j_M}\mathbf{L}^{i'_1\ldots i'_N}_{\mathbf{y}j'_1\ldots j'_M}c_{i_1\ldots i_N i'_1\ldots i'_N}(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$
3. *The following two conditions hold simultaneously:*
   - $\mathbf{L}^{i_1\ldots i_N}_{\mathbf{x}j_1\ldots j_M}E[f_{i_1\ldots i_N}(\mathbf{x})] = \mathbf{k}_{j_1\ldots j_M}(\mathbf{x})$
   - $\mathbf{L}^{i_1\ldots i_N}_{\mathbf{x}j_1\ldots j_M}c_{i_1\ldots i_N i'_1\ldots i'_N}(\mathbf{x},\mathbf{y}) = 0 \forall \mathbf{y}$

4.1 Simple Example: Heat Flow

A simple physical information about heat flow S^Q in the upper continental crust is that it is (nearly) source free:

$$\nabla\mathbf{S}^Q \approx 0 \tag{12}$$

We need to consider sources at the lower heating surface and sinks at the upper cooling surface. The heat flow in an scenario with equal heat flow

over the surfaces can be used as trend $\mathbf{S}_0^Q(\mathbf{x}$. The variogram, which is necessarily instationary in depth direction, could be estimated from computer simulations of the heat flow through the radom heterogenous ore bodies or from measured data. This differential equation does not fully describe the process, even when all boundary conditions are known. However this is not necessary for the application of kriging. Based on the obtained variogram we can estimate the heat flow at unmeasured locations by the simple method of kriging without solving the full nonlinear problem. Although not based on exact physics the estimation preserves the energy in the system, when the variogram is admissible for equation 12.

4.2 Special Case: Spatial Phenomena Related by Differential Equations

Besides the implication of this theorem for covariograms and kriging of tensors we get a simple application, when differential equations link the spatial behavior of more than one physical quantity. A simple example is a potential field with diffuse sources randomly distributed. Denotes $\phi(\mathbf{x})$ the potential and $q(\mathbf{x})$ the intensity field of sources we normally have the Poisson equation:

$$\Delta\phi = q \tag{13}$$

The Poisson equation has often been considered for geostatistical applications (e.g. [Chilès&Delfiner 1999]). Examples are the electric potential ϕ with electric loading density q and the gravitation potential ϕ with mass density q. Let us think of ϕ and q as random quantities, we can see this as a differential equation of the bivariate random field

$$Z(\mathbf{x}) = \begin{pmatrix} \phi(\mathbf{x}) \\ q(\mathbf{x}) \end{pmatrix}$$

$$\begin{pmatrix} \Delta & -1 \end{pmatrix} \begin{pmatrix} \phi(\mathbf{x}) \\ q(\mathbf{x}) \end{pmatrix} = 0$$

From corollary 8 we get for the tensorial covariogram

$$\mathbf{c}(\mathbf{x},\mathbf{y}) := \begin{pmatrix} c_{\phi\phi}(\mathbf{x},\mathbf{y}) & c_{\phi q}(\mathbf{x},\mathbf{y}) \\ c_{q\phi}(\mathbf{x},\mathbf{y}) & c_{qq}(\mathbf{x},\mathbf{y}) \end{pmatrix}$$

the differential equation

$$\begin{pmatrix} \Delta_{\mathbf{x}} & -1 \end{pmatrix} \mathbf{c}(\mathbf{x},\mathbf{y}) = 0 \forall \mathbf{y}$$

or equivalently

$$\left(\begin{array}{cc} \Delta_{\mathbf{y}} & -1 \end{array} \right) \left(\left(\begin{array}{cc} \Delta_{\mathbf{x}} & -1 \end{array} \right) \mathbf{c}(\mathbf{x},\mathbf{y}) \right)^t |_{\mathbf{x}=\mathbf{y}} = 0$$

And thus a condition on the relation of the covariograms of the two quantities:

$$\Delta_{\mathbf{x}}\Delta_{\mathbf{y}} c_{\phi\phi}(\mathbf{x},\mathbf{y}) - 2\Delta_{\mathbf{x}} c_{\phi q}(\mathbf{x},\mathbf{y}) + c_{qq}(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$$

5 Ideas to Construct Physically Admissible Variograms

An important problem is to construct variograms and covariograms which solve the equation

$$L_{\mathbf{x}} L_{\mathbf{y}} c(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{y}} = 0$$

I can give four methods, too old to be found in the books I read.

5.1 Construction by Superposition

Conic combinations of variograms and covariograms are itself variograms and covariograms[Christakos 1992].

$$c(\mathbf{h}) := \sum_{i=1}^{q} w_i c_i(\mathbf{h}), \quad w_i \geq 0$$

Theorems 5 and 3 reduce the condition of the differential equation to some linear conditions on the low order derivatives of $c(h)$ or $c(\mathbf{h})$. They could be used to construct admissible covariograms or variograms, since the derivatives of $c(\mathbf{h})$ are linear combinations of the derivatives of the $c_i(\mathbf{h})$, by solving the set of linear equations in w_i given by:

$$\sum_{i=1}^{q} w_i (L_{\mathbf{h}} L_{\mathbf{h}} c_i)(0) = 0$$

with positive coefficients w_i

5.2 Construction from known solutions

Suppose we have for every $\mathbf{z}$ a $w(\mathbf{z},\mathbf{x})$ that is a solution of

$$L_x w(\mathbf{z},\mathbf{x}) = 0$$

which is uniformly square integrable:

$$\int w(\mathbf{z},\mathbf{x})^2 d\mathbf{z} < a \in \mathbb{R} \forall \mathbf{x} \tag{14}$$

E.g. when the the equation is stationary and there is a solution $w(\mathbf{h})$ with finite support we can write:

$$w(\mathbf{z},\mathbf{x}) := w(\mathbf{x}-\mathbf{z})$$

Then

$$c(\mathbf{x},\mathbf{y}) := \int w(\mathbf{z},\mathbf{x}) w(\mathbf{z},\mathbf{y}) d\mathbf{z}$$

is a positive semidefinite function which is admissible for the differential equation. c is semidefinite since for any square integrable function f we get

$$\begin{aligned} & \int\int f(\mathbf{x}) c(\mathbf{x},\mathbf{y}) f(\mathbf{y}) d\mathbf{x} d\mathbf{y} = \\ = & \int\int f(\mathbf{x}) w(\mathbf{z},\mathbf{x}) d\mathbf{x} \int w(\mathbf{z},\mathbf{y}) f(\mathbf{y}) d\mathbf{y} d\mathbf{z} = \\ = & \int \left(\int f(\mathbf{x}) w(\mathbf{z},\mathbf{x}) d\mathbf{x} \right)^2 d\mathbf{z} \end{aligned}$$

The integrals can be permuted according to Fubini since all integrals exist due to eq. (14).

c solves the differential equation since:

$$\begin{aligned} L_{\mathbf{x}} c(\mathbf{x},\mathbf{y}) &= L_{\mathbf{x}} \int w(\mathbf{z},\mathbf{x}) w(\mathbf{z},\mathbf{y}) d\mathbf{z} = \\ &= \int L_{\mathbf{x}} w(\mathbf{z},\mathbf{x}) w(\mathbf{z},\mathbf{y}) d\mathbf{z} = \\ &= \int 0 w(\mathbf{z},\mathbf{y}) d\mathbf{z} = 0 \end{aligned}$$

The differential and the integration permute according to the theorem of Fubini, since the inner integrals exist.

5.3 Construction Using Spectral Densities

Every valid stationary covariogram can be written [Cressie 1993]

$$c(\mathbf{h}) = \int cos(\omega^t \mathbf{h}) dG(\omega)$$

with a spectral measure G on $\mathbb{R}^d$. Its even derivatives are given by

$$
\begin{aligned}
& \frac{d^i}{d\mathbf{h^i}} c(\mathbf{h})|_{\mathbf{h}=0} = \\
= & \int \frac{d^i}{d\mathbf{h^i}} cos(\omega^t \mathbf{h}) dG(\omega)|_{\mathbf{h}=0} = \\
= & \int -1^{i/2} \left(\otimes_{j=1}^{i} \omega\right) \underbrace{cos(\omega^t \mathbf{h})}_{1} dG(\omega)|_{\mathbf{h}=0} = \\
= & -1^{i/2} E\left[\otimes_{j=1}^{i} \omega\right] dG(\omega) =: \mu^{(i)}
\end{aligned}
$$

Thus the even derivatives (cmp. [Christakos 1992, p. 52 remark 4]) are simply the even non centered moments of the spectral measure, and thus the differential equation

$$L_{\mathbf{h}} L_{\mathbf{h}} c(\mathbf{h})|_{\mathbf{h}=0} = 0$$

with

$$L_{\mathbf{h}} L_{\mathbf{h}} = \sum_{i=0}^{2n} \beta^i \frac{d^i}{d\mathbf{h}^i}$$

is transformed to

$$\sum_{i=0}^{2n} \beta_{pj_1 \ldots j_i}^{(i)} \mu_{j_1 \ldots j_i}^{(i)} = 0 \forall p$$

which is a linear condition on the even moments of G.

Remark 2 *Since all even fourth order moments are positive definite, no nontrivial stationary covariogram is admissible for $\Delta f(x) = 0$, which is valid in source free areas of a potential field. Thus e.g. the Gaussian variogram or any other stationary stationary variogram is inadmissible for modeling the gravity potential of the earth in free space. We must use nonstationary covariogram models.*

5.4 Modeling Variograms from Relations

The problem to determine the variogram of the gravity potential of the earth leads to a different construction (compare [Chilès&Delfiner 1999]). Let us reconsider the example eq. (13) and use the slightly more general from

$$L_{\mathbf{x}} \phi = q$$

Possibly we have a (nonstationary) model for the covariogram $c_{qq}(\mathbf{x}, \mathbf{y})$ of q from other investigations such as seismic tomography of the earth. Then using a Green function g, which is the solution of

$$L_{\mathbf{x}} g(\mathbf{x}) = \delta_{\mathbf{x}}$$

where $\delta_{\mathbf{x}}$ is the delta distribution in $x = 0$. Then we get as solution a valid covariogram by double convolution of $c_{qq}(\mathbf{x}, \mathbf{y})$ with $\begin{pmatrix} g(\mathbf{x}), & \delta_{\mathbf{x}} \end{pmatrix}$

$$\mathbf{c}(\mathbf{x}, \mathbf{y}) = \int_{\mathbf{x}'} \int_{\mathbf{y}'} \begin{pmatrix} g(\mathbf{x}) \\ \delta_{\mathbf{x}} \end{pmatrix} c_{qq}(\mathbf{x}', \mathbf{y}') \begin{pmatrix} g(\mathbf{x}) & \delta_{\mathbf{x}} \end{pmatrix} d\mathbf{x}' d\mathbf{y}'$$

In the situation of the gravity field of the earth the Green function is

$$g(\mathbf{x}) = \frac{g_0}{\|x\|}$$

where g_0 is the general gravity constant.

6 Variograms and Generalized Covariograms

All the theory has been formulated in terms of the covariogram. However in the application of kriging most often a trend model is used, which implies the usage of a variogram or a generalized variogram or generalized covariogram in the context of IRFk[Chilès&Delfiner 1999][Cressie 1993]. In these situations we have an underlying trend model:

$$E[f(\mathbf{x})] = f_0(\mathbf{x}) + \sum_{j=1}^{p} \beta_j f_j(\mathbf{x}) = f_0(\mathbf{x}) + \beta^t \begin{pmatrix} f_j(\mathbf{x}) \\ \vdots \\ f_p(\mathbf{x}) \end{pmatrix}$$

In general it only makes sense to have:

$$L_{\mathbf{x}} f_j = 0, \ i \neq 0 \text{ and } \quad L_{\mathbf{x}} f_0 = k(\mathbf{x})$$

since otherwise only a subspace of choices for β would solve the differential equation.

When we use a trend model of this type the generalized covariogram and the generalized variogram can be inferred and used only up to equivalence relation given by[Chilès&Delfiner 1999]:

$$c(\mathbf{x}, \mathbf{y}) \equiv c'(\mathbf{x}, \mathbf{y}) :\Leftrightarrow c(\mathbf{x}, \mathbf{y}) - c'(\mathbf{x}, \mathbf{y}) \in \langle \{ f_i(\mathbf{x}) f_j(\mathbf{y}) : i, j = 1, \ldots, p \} \rangle$$

or more precisely for stationary covariograms and stationary trend models:

$$c(\mathbf{h}) \equiv c'(\mathbf{h}) :\Leftrightarrow c(\mathbf{h}) - c'(\mathbf{h}) \in \left\langle \left\{ f_i(\mathbf{x}) f_j(\mathbf{x} + \mathbf{h}) : i, j = 1, \ldots, p, \mathbf{x} \in \mathbb{R}^d \right\} \right\rangle$$

This is a slight generalization of the theory of intrinsic random functions, where only the even part of this equivalence relation is used since generalized

stationary covariograms are always even; i.e. they can be restricted to even functions. The $\langle\ \rangle$ brackets denote the linearly generated function space.

When we consider the variogram

$$2\gamma(\mathbf{x},\mathbf{y}) = c(\mathbf{x},\mathbf{x}) + c(\mathbf{y},\mathbf{y}) - 2c(\mathbf{x},\mathbf{y})$$

we realize, that with $L_{\mathbf{h}}1 = 0$ it solves the differential equation

$$0 = L_{\mathbf{y}}L_{\mathbf{x}}\gamma(\mathbf{x},\mathbf{y}) = \frac{1}{2}L_{\mathbf{x}} \underbrace{L_{\mathbf{y}}c(\mathbf{x},\mathbf{x})}_{0,\text{since } L_{\mathbf{y}}1=0} + \frac{1}{2}L_{\mathbf{y}} \underbrace{L_{\mathbf{x}}c(\mathbf{y},\mathbf{y})}_{0,\text{since } L_{\mathbf{x}}1=0} - \underbrace{L_{\mathbf{y}}L_{\mathbf{x}}c(\mathbf{x},\mathbf{y})}_{0} \tag{15}$$

if and only if $c(\mathbf{h})$ solves it. And it solves the differential equation

$$0 = L_{\mathbf{y}}L_{\mathbf{x}}\gamma(\mathbf{x},\mathbf{y}) = \frac{1}{2} \underbrace{L_{\mathbf{x}}c(\mathbf{x},\mathbf{x})}_{2*L_{\mathbf{x}}c(\mathbf{x},\mathbf{y})|_{\mathbf{y}=\mathbf{x}}=0,\text{since } L_{\mathbf{y}}1=0} + \frac{1}{2} \underbrace{L_{\mathbf{x}}c(\mathbf{y},\mathbf{y})}_{0,\text{since } L_{\mathbf{x}}1=0} - \underbrace{L_{\mathbf{x}}c(\mathbf{x},\mathbf{y})}_{0} \tag{16}$$

Thus if the trend model may contain an unknown mean, due to $L_{\mathbf{x}}1 = 0$, then the restrictions on the covariogram equivalently apply to the variogram.

With the same arguments we get the general theorem for generalized variograms and covariograms:

Theorem 9 *If*

$$E[f(\mathbf{x})] = f_0(\mathbf{x}) + \sum_{j=1}^{p} \beta_j f_j(\mathbf{x}) = f_0(\mathbf{x}) + \beta^t \begin{pmatrix} f_j(\mathbf{x}) \\ \vdots \\ f_p(\mathbf{x}) \end{pmatrix}$$

is a valid trend model for the differential equation

$$L_{\mathbf{x}}f(\mathbf{x}) = k(\mathbf{x})$$

i.e.

$$L_{\mathbf{x}}f_j = 0,\ i \neq 0 \text{ and } \quad L_{\mathbf{x}}f_0 = k(\mathbf{x})$$

Then the consequences of theorems 1, 2, 3, 5, 6, 7, 8 stay unchanged when we replace the covariogram with a generalized covariogram or generalized variogram.

Proof: It is sufficient to proof theorem 1 for generalized covariograms and generalized variograms, since all the other theorems follow from it. Thus we need a modified proof for theorem 1. The definition of a generalized covariance, as given in [Chilès&Delfiner 1999, p. 253], implies that it yields

the covariance for any linear operator, which is a finite linear combination of evaluated measurements

$$L_{\mathbf{x}} f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i f(\mathbf{x}_i)$$

that maps the trend to zero:

$$L_{\mathbf{x}} f_i(\mathbf{x}) = 0,\ i = 1, \ldots, p \tag{17}$$

The linear equations (17) define a linear subspace of all linear operators called Λ_x. The special $L_{\mathbf{x}}$, which is used in the assumptions of theorem 9 is in $\Lambda_{\mathbf{x}}$ by definition. Thus we just need to write $L_{\mathbf{x}}$ as a limit of finite linear combinations in Λ_x. This is possible, since the linear combinations are dense in that subspace, since derivatives are just members in the closing of finite linear combinations. The same holds for the generalized variogram, since it is related to the corresponding covariogram by a factor -1, which does not alter the solution of a linear differential equation.

7 Conclusion

The combination of structural knowledge given by differential equations and empirical knowledge given by observations can be handled by without any involved mathematics by kriging based on partial differential equations. The differential equations essentially impose a restriction on the covariogram or variogram. Using these admissible variograms or covariograms the kriging results solve the differential equations. Thus this method combines the advantage of kriging to use probabilistic information from estimated variograms with the advantage of numerical methods using the physical knowledge about the process. Thus kriging provides a simple solution to a big class of problems.

References

[Chilès&Delfiner 1999] Chilès, J.-P., Delfiner, P., 1999, *Geostatistics: Modelling Spatial Uncertainty*, J. Wiley& Sons, New York

[Christakos 1992] Christakos, G. (1992) *Random Fields Models in Earth Sciences*, Academic Press, Inc., San Diego

[Cressie 1993] Cressie, N.A.C., 1993, *Statistics for Spatial Data (rev. ed.)*: J. Wiley & Sons, New York

[Gard 1988] Gard, T. C. (1988) *Introduction to stochastic differential equations*, Monographs and textbooks in pure and applied mathematics, Marcel Dekker Inc., New York

[Müller 1998] Müller, Claus (1998), *Analysis of Spherical Symmetries in Euclidian Spaces*, Springer, New York

[Wackernagel 1998] Wackernagel, H. (1998), *Multivariate Geostatistics: An Introduction with Applications*, second completely revised edition, Springer, Berlin