4079

On the choice of norm and base functions for the solution of the inverse gravimetric problem.

by

G.Hein, University FAF, Muenchen, FRG,
F.Sanso', Politecnico di Milano, Milano, Italy,
G.Strykowski and C.C.Tscherning,
Geodetic Institute, Charlottenlund, Denmark.

Abstract:

The inverse gravimetric problem will have a unique solution, if the density distribution is required to be an element of a suitable Hilbert space, and a condition of minimum norm is applied. Examples of suitable spaces are spaces spanned by harmonic density distributions or by linear combinations of disjoint indicator functions, in both cases using Sobolev-type inner products. The space will also have a reproducing kernel, which may be regarded as an implicitly imposed covariance kernel for the evaluation functionals, thereby also defining an auto and cross-covariance function for and with the gravity anomalies.

A comparison with empirically estimated covariance functions for density (anomaly) and gravity anomaly data close to the Rhinegraben area, FRG, shows that the above mentioned spaces impose either a too strong or a too weak covariance between the quantities. Therefore, the use of other alternatives becomes necessary. Here indicator functions with overlapping support seems to be realistic, and flexible enough to provide the proper covariance pattern.

Revised version of a paper presented at 17. Conference on Mathematical Geophysics, Blanes, Cataluna, Spain, June 1988.

1. Introduction

Traditionally, the inverse gravimetric problem is defined as the problem of determining the (mass) density distribution of the Earth from gravity data measured at the surface of the Earth. We will, however, here consider the more general problem of determining the density distribution from all kinds of data, including in-situ density observations, geoid undulations obtained from satellite radar altimetry, and ultimately seismic data. Hereby the solution of the inverse gravimetric problem will contribute to the improved modelling of the exterior gravity field, which is one of the important tasks of geodesy.

The mathematical formulation of the general problem therefore includes the selection of one or several funtion spaces (see e.g. Backus and Gilbert, 1967). We will here only consider the selection of the function space, which contain elements which can be expected to be good approximations to the density distribution. Furthermore, since we will only consider linearized observation functionals, we will only deal with the density contrast, or density anomaly function, d. This function is equal to the difference between a long wavelength, primarily radial dependent, density reference function, ρ_o and the true density, ρ, i.e. $d = \rho - \rho_o$. The potential of d is the anomalous potential, T,

$$T(P) = G \cdot \int_{\Omega} \frac{d(Q)}{|P-Q|} \, d\Omega_Q \tag{1}$$

where Ω is the Earth, Q a point inside the Earth, G the gravity constant and P an arbitrary point.

A function space which has the structure necessary for the solution of the inverse gravimetric problem is prescribed through the base functions, which span the space, and an inner product. This may be done in many different ways, but two principles may serve as guidelines:

(a) the mathematical and computational simplicity of the function space

(b) the physical (geological) properties which are modelled by the space.

In (b) we will also include the (not necessarily complete) modelling of statistical information like density variation and spatial correlation. This information is sometimes labelled "a-priori" information (see e.g. Jackson, 1979).

In our age of the computer, principle (a) may be a good starting point. The computer may be used as an experimental tool, and if one is so lucky as to find a mathematically simple model, which agrees with the physical reality, then there should be a good chance that the inverse problem has a reasonable solution.

We will here report on a numerical experiment, which however did not lead us to a realistic solution of the inverse problem. But it gave us some further insight.

In section 2 we describe the numerical experiment, where we used anomalous density distributions, which were harmonic, and selected an inner product based on statistical information. Since in performing the experiment it became clear that harmonicity is too strong a condition, we looked into the possibility of using the traditional indicator functions. These functions generally have disjoint support, thereby making it impossible to prescribe a simple inner product leading to correlated density values, except 100 % or 0 % correlations. In section 3 we propose to use indicator functions with a certain joint support. This leads to correlated density values.

2. Harmonic inversion of gravity data

The general inverse gravimetric problem is a typical approximation problem in a function space. We have a finite number of observations, and the function we want to determine is an element of an infinite dimensional space. The brute force way to find the function is to select a subspace of dimension less than or equal to the number of observations, and then determine a unique element using a least squares or mimimum norm condition. However, this selection may be quite arbitrary, and it seems better to build a more general model first. Such a model is e.g. the traditional set of disjoint indicator functions, by which a geological structure is broken into pieces. Another alternative is to use so-called quasi-harmonic functions, i.e. functions which are nearly harmonic,

$$\Delta(d/f(r)) = 0 \tag{2}$$

where Δ is the Laplace operator and f a non-zero function of the radial distance, r. If f(r) = 1, d is harmonic. The advantage of using these functions is described in (Tscherning and Suenkel, 1981) and (Tscherning and Strykowski, 1987). We will here only consider harmonic d, since it for local applications seems of limited importance how f is selected. (We have tried various $f(r) = r^m$, m an integer).Then (from now on using spherical approximation)

$$d(\phi,\lambda,r) = \sum_{i=2}^{\infty} \left(\frac{r}{R}\right)^i \sum_{j=-i}^{i} c_{ij} Y_{ij}(\phi,\lambda), \tag{3}$$

with

$$Y_{ij}(\phi,\lambda) = \bar{P}_{ij}(\sin\phi) \begin{cases} \cos j\lambda & i \geq j \geq 0 \\ \sin|j|\lambda & 0 > j \geq -i \end{cases}$$

where ϕ is the latitude, λ the longitude, R the mean radius of the Earth, $\bar{P}_{ij}$ the normalized associated Legendre functions, and c_{ij} constants.

This space may be equipped with an inner product, so that the space is endowed with a reproducing kernel, K(P,Q),

$$K(P,Q) = \sum_{i=2}^{\infty} \sum_{j=-i}^{i} \left(\frac{rr'}{R^2}\right)^i \sigma_{ij} Y_{ij}(\phi,\lambda) Y_{ij}(\phi',\lambda'), \tag{4}$$

where σ_{ij} are positive constants and (ϕ',λ',r') are the spherical coordinates of Q. If the constants are selected so that they are independent of j and for each i equal to the mean square sum of the coefficients c_{ij} in (3), then the space may be used to obtain an optimal estimate of d. It is optimal in a least squares sense, described in Sanso' (1986). The reproducing kernel will then have the interpretation of an implicitly imposed auto-covariance function of the density anomalies.

The constants c_{ij} in (3) are naturally not known. However, the corresponding square sum of the coefficients of the anomalous potential T may be estimated on parametric form, see e.g. (Knudsen, 1987).

With

$$\sigma_i(d) = \sum_{j=-i}^{i} c_{ij}^2$$

the corresponding "degree-variances" of T becomes using (1)

$$\sigma_i(T) = \sigma_i(d) \left(\frac{4\pi R}{(2i+1)(2i+3)}\right)^2 \quad (5)$$

The parameter form used for $\sigma_i(T)$ is

$$\sigma_i(T) = a \cdot \varepsilon_i \quad , \; i < m$$
$$\sigma_i(T) = A/((i-1)i-2)(i+B)) \cdot \left(\frac{R_B}{R}\right)^{2i+2} \; , \; i \geq m \quad (6)$$

where A and R_B are real constants ($R_B < R$) and B is an integer constant. This permit the numerical evaluation of the reproducing kernel using closed expressions (Tscherning and Rapp, 1974). The constants $a.\varepsilon_i$ describes the variation due to regional residual effects caused by the uncertainty in the reference density function ρ_o or corresponding potential, W_o.

The use of models like the one given by eq. (6) corresponds to the use of a weighted Sobolev-type inner product, see Tscherning (1972,1986).

Having fixed the reproducing kernel (4), the inverse problem is readily solved using the method of collocation, i.e. with observations $x_i = L_i(d) + v_i$,

$$\tilde{d}(P) = \{K(P,L_i)\}^T \{K(L_i,L_j) + \tau_{ij}\}^{-1} \{x_j\} \qquad (7)$$

where $K(L_i,L_i)$ is equal to the observation functionals L_i,L_j applied on K, v_i the observation error with variance-covariance τ_{ij}, and d(P) is the estimated density value in P.

The software for estimating the covariance function parameters, and for the evaluation of (7) is described in Forsberg et al. (1988) and Hein et al. (1987).

The harmonic inversion method was used on gravity and (surface) density data from an area close to the Rhine Graben, FRB, see Fig. 1. Totally 104 gravity and density values were available in an area of externt 25 km x 25 km. The small area makes the covariance function parameter estimation difficult. However, the auto-covariance function of the gravity anomalies was modelles with a reasonable success, see Fig. 2. But the simultaneous modelling of the auto-covariance function and the cross-covariance function between gravity and density values was not successful, see Fig. 3 and 4. It was impossible to obtain a density-gravity cross-correlation as small as the one observed, and still keeping the observed gravity anomaly variance. The model cross-correlation was generally around 90 %, while the observed value was around 65 %.

The reason for this, probably, is the harmonicity of the density functions. This property will always force the maximal variance to be at the boundary surface. It also causes the strong cross correlation between the gravity and the density values. We did, however, anyway carry out a density estimation using eq. (7) with gravity anomalies as observed values. The result is shown in Fig. 5. Gravity were also predicted from density values, and from combined gravity and density values. The results were not satisfactory, in the sense that the standard deviation of the differences between observed and computed quantities was close to the standard deviation of the observed values.

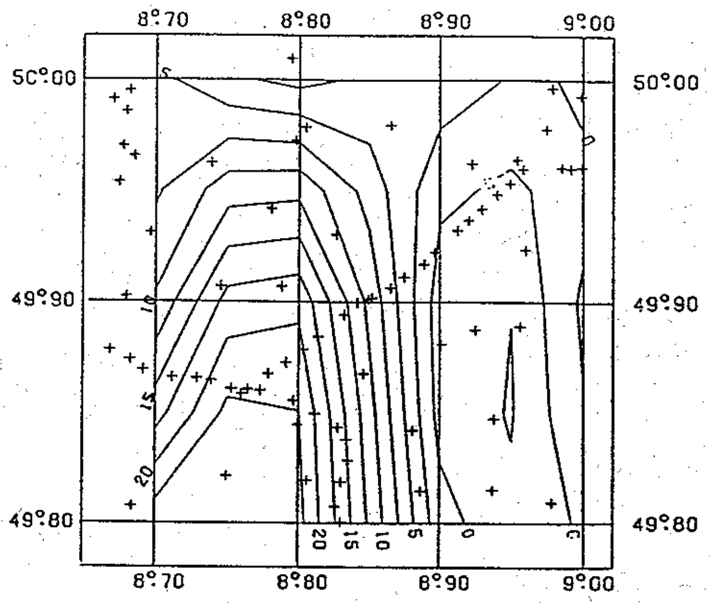


Fig. 1a. The position of the stations in the area and the bouguer anomalies (in mgal).

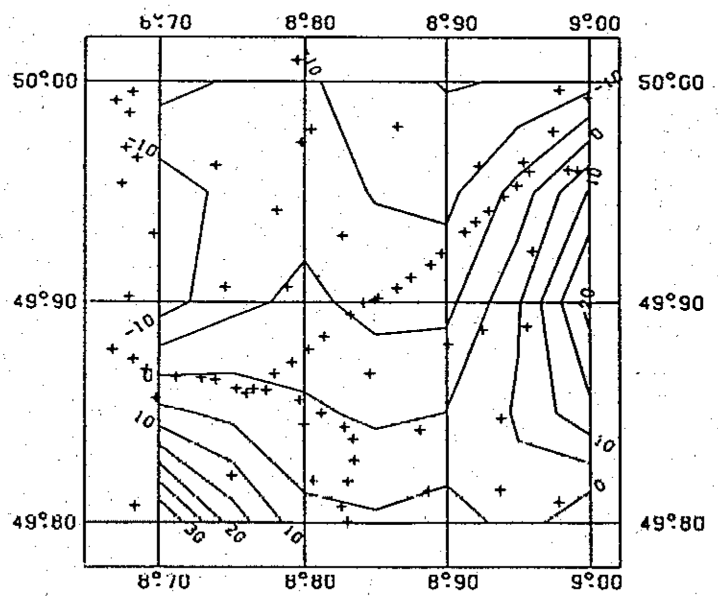


Fig. 1b. The position of the stations and the surface values of density anomalies (the absolute density - 2.67 gcm^{-3}). The units of density are 0.01 gcm^{-3}.

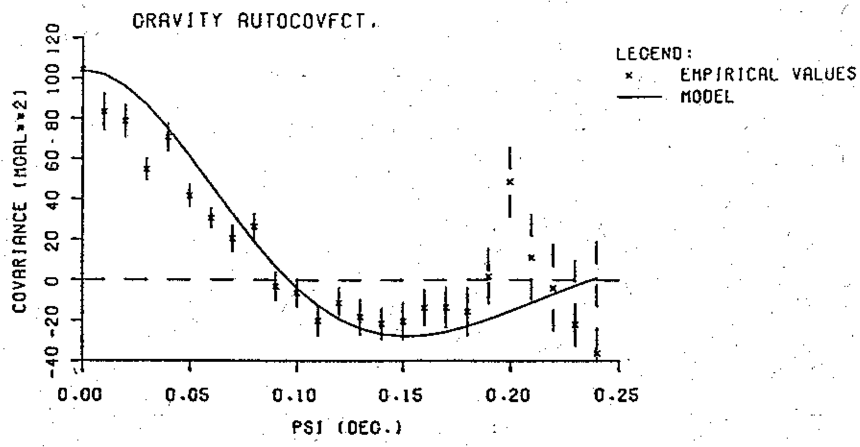


Fig. 2. The empirical- and model- auto-covariance functions for gravity anomalies at the surface ($R-R_B$ = 6000 m, m = 1100, $C_{\Delta g \Delta g}(0)$ = 104.00 $mgal^2$).

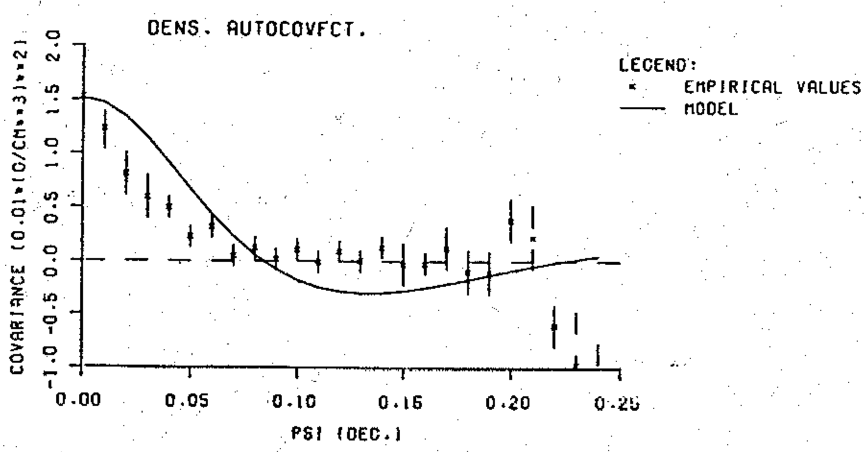


Fig. 3. The empirical- and model auto-covariance functions for density anomalies at the surface.

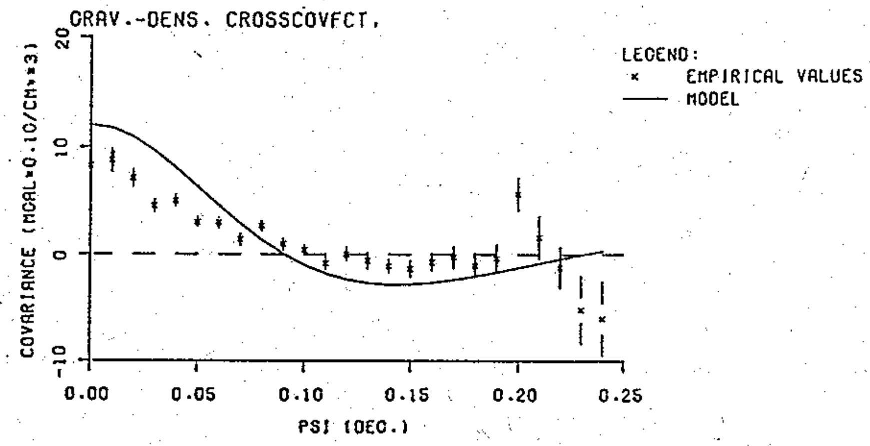


Fig. 4. The empirical- and model- cross-covariance function for the surface- gravity- and density- anomalies.

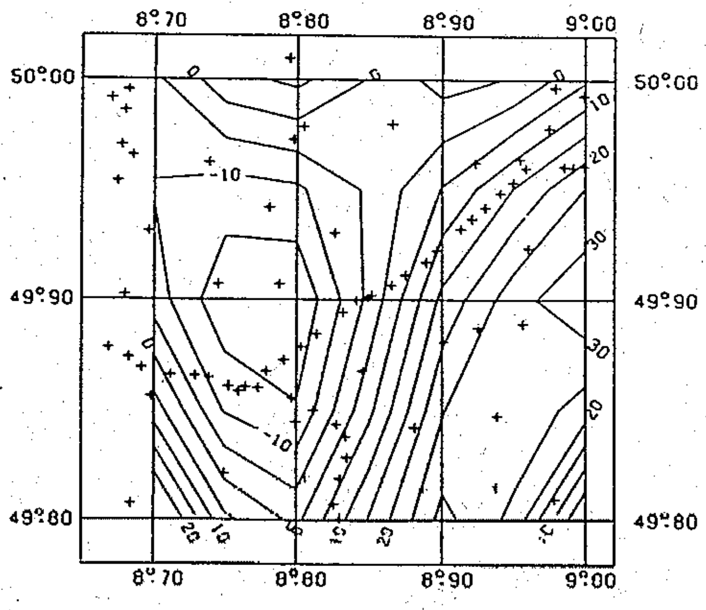


Fig.5. The differences between the predicted- and the measured density anomalies in units of 0.01 gcm^{-3}. The parameters of the model are : ($R-R_B$ = 5500 m, m = 1100, $C_{\Delta g \Delta g}(0)$ = 104.00 $mgal^2$).

This negative result is, however, not completely negative. At least we have verified that we have a method to solve the inverse gravimetric problem.

3. Use of base functions with overlapping support

Since the use of harmonic base functions - at least for the used data set - caused difficulties, other alternatives must be considered. However, we must be sure that the density covariance function implied by the selected function space can be used to model the empirically observed function. Also, it would be important to be able to use information on lateral density discontinuities. (Radial discontinuities may be modelled using a discontinuous function f in eq. (2)).

The most frequently used set of base functions is the set of indicator functions with disjoint support. They are very flexible, and may represent a geological structure nicely. In (Sanso' and Tscherning, 1982) it is proposed to use such a function to model the density down to a certain sphere, and then use harmonic density distributions inside the sphere. The inner product must in this case be the one related to the L_2-norm, used for the indicator functions, and a Sobolev type norm for the harmonic part,

$$|d|^2 = \int_{\Omega_o} d^2 d\Omega_o + \int_{\Omega_1} (F(d))^2 d\Omega_1 \qquad (8)$$

where Ω_o and Ω_1 are the sets outside and inside a sphere, respectively, and F is a suitable differential operator.

However, this choise implicitly corresponds to the use of a density covariance function at the Earth's surface with only values equal to one and zero,

$$K(P,Q) = \sum_{i=1}^{n} I_i(P)I_i(Q)/v_i, \qquad (9)$$

where I_i is the indicator function of the i'th block and v_i is the volume of the block. (This is because the normalized base functions are

$$\bar{I}_i(P) = I_i(P)/(v_i)^{1/2}).$$

One way to overcome this problem is to adopt a covariance function for the points, dependent on the distance between the points, for example. This may be possible working in 2 dimensions, but for 3-dimensional modelling it seems very difficult to find analytic covariance models different from these described in section 2. The main problem being, that we must be able to calculate $L_i K(P,Q)$, where L_i for example is the Newton functional, eq. (1).

A simple, and computationally easy, way to prescribe a covariance function with values different from 1 and 0, is to use overlapping indicator functions.

This is clearly illustrated using 3 blocks, each with volume equal to 1, and having a 50 % overlap, see Fig. 6.

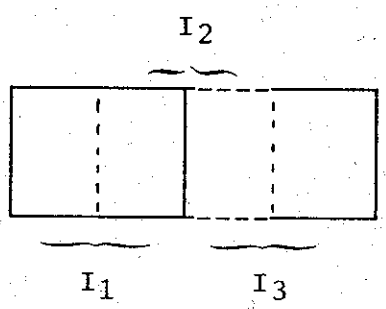


Fig. 6. Overlapping base function.

The corresponding reproducing kernel (eq. (9)) is computed in Appendix 1. The resulting correlation function is shown in Fig. 7, as a function of distance between the points P and Q.

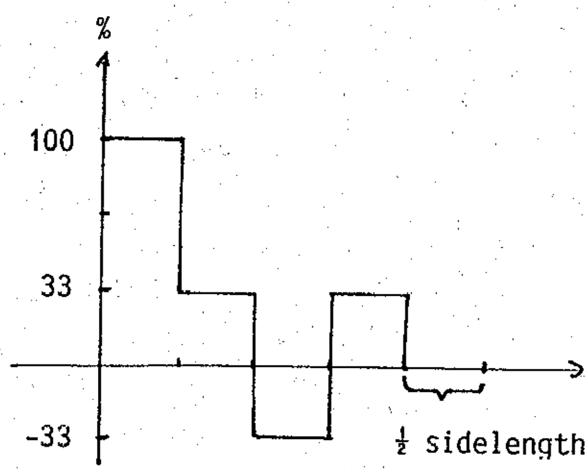


Fig. 7. Correlation function with 3 blocks.

The calculation of the value of linear functionals applied on the reproducing kernels is most easily done without the explicit calculation of the expansion of the kernel in the old base functions. Since Cholesky decomposition of a positive definite matrix and the Gram-Smith orthonormalization process are equivalent operations, we may take advantage of this.

Suppose the matrix C as elements have the inner products of all the base functions,

$$C_{ij} = \langle I_i, I_j \rangle,$$

and that C has the Cholesky decomposition

$$C = U^T U,$$

where U is an upper triangular matrix. Then

$$\{\bar{I}_j\} = \{U_{ij}\}^{-1} \{I_j\}$$

is the vector of orthonormalized base functions. This is easily seen, because of the linearity of the inner product

$$\langle \{\bar{I}_j\}, \{\bar{I}_k\} \rangle = \langle U^{-1}\{I_j\}, U^{-1}\{I_k\} \rangle =$$

$$(U^{-1})^T \langle \{I_j\}, \{I_k\} \rangle U^{-1}) = (U^{-1})^T C U^{-1} = Id,$$

where Id is the identity matrix.

Consider now, for example, the gravity disturbance functional $\delta g(P)$ evaluated in P, and let $\delta g_i(P)$ be the gravity in P caused by the i'th indicator function,

$$\delta g_i(P) = G \int_\Omega I_i(Q) \frac{\partial}{\partial r} \left(\frac{1}{|P-Q|}\right) d\Omega = L_{\delta g(P)}(I_i).$$

Then the implied covariance between two gravity disturbance values becomes

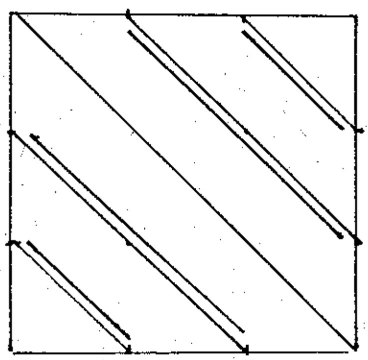
$$L_{\delta g(P)}(L_{\delta g(Q)}K(P,Q)) =$$

$$= \sum_{i=1}^{n} L_{\delta g(P)}(\bar{I}_i)L_{\delta g(Q)}(\bar{I}_i)$$

$$= (U^{-1}\{\delta g_i(P)\})^T (U^{-1}\{\delta g_i(Q)\}) = K(\delta g(P), \delta g(Q)).$$

These considerations also show (the proof is left as an exercise to the reader) that the use of correlated density values corresponds to the introduction of unknown density values for each block as parameters in collocation. However, the parameters are correlated with C as the variance-covariance matrix.

It also shows the problems associated with the use of indicator functions, namely the rather large computational task. However, the C matrix will be sparse, since generally only a small number of blocks will overlap. If the block structure is selected in a regular manner, then sparse matrix techniques may be applied, see Fig. 8.



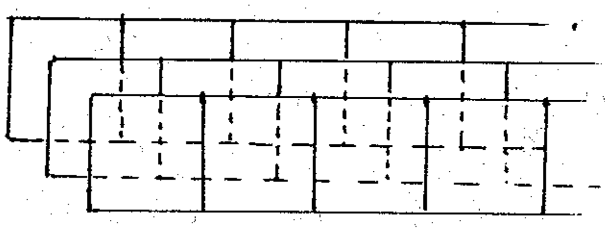


Fig. 8. Non-zero elements in C matrix, where 3 blocks overlap.

4. Conclusion

The use of a combination of quasi-harmonic functions and base functions with overlapping support should give us a realistic and flexible tool to provide a proper covariance pattern. Numerical investigations are in progress, but have not yet been completed.

A considerable number of new and quite simple mathematical models are now available in order to study the inverse gravimetric problem. However, other slightly more complicated models, using indicator functions multiplied by linear functions in the (Cartesian) coordinates (Sanso'et al., 1986) should also be investigated. They have the theoretical advantage that more general inner products involving e.g. the derivatives of the functions could be used. Since the solution of the inverse problem in this case corresponds to the quadratic minimalization of density derivatives, it might lead to density estimates with interesting physical properties.

The problem in future investigations will not so much be the lack of mathematical models, as the lack of suitable sets of test data. Where do we find sets of density values distributed well, both with respect to depth and horizontally? The data sample from the area close to the Rhine Graben is unique, but more such samples from regions with different geological characteristics, should be made available.

Acknowledgements: This investigation has in part been sponsored by Norsk Hydro Udforskning.

Appendix. The reproducing kernel of a Hilbert space spanned by 2 and 3 indicator functions with 50 % overlap.

The covariance function is the product sum of the orthonormalized, (Gram-Smith), functions:

$$|I_1|^2 = 1, \text{ i.e. } \quad \bar{I}_1 = I_1,$$

where the bar indicates normalization. Since

$$\langle \bar{I}_1, I_2 \rangle = \frac{1}{2} \text{ , we have}$$

$$\bar{I}_2 = (I_2 - \langle \bar{I}_1, I_2\rangle \ \bar{I}_1)/(|I_2 - \langle \bar{I}_1, I_2\rangle \ \bar{I}_1|)$$

$$= (I_2 - \frac{1}{2}\ \bar{I}_1)/(\frac{3}{4})^{1/2} = \frac{2}{\sqrt{3}}\ (\bar{I}_2 - \frac{1}{2}\ I\ )\ .$$

This means, that with only two functions we have

$$K_2(P,Q) = \bar{I}_1(P)\ \bar{I}_1(Q) + \bar{I}_2(P)\ \bar{I}_2(Q)$$

$$= I_1(P)\ I_1(Q) + \frac{4}{3}\ (I_2(P)\ I_2(Q) + \frac{1}{4}\ I_1(P)\ I_1(Q)$$

$$- \frac{1}{2}\ \{I_1(P)\ I_2(Q) + I_1(Q) I_2(P)\}$$

$$= \frac{4}{3}(I_1(P)I_1(Q)+I_2(P)I_2(Q) - \frac{1}{2}\ (I_1(P)I_2(Q)+I_1(Q)I_2(P))). \quad (10)$$

The evaluation functionals, Ev_P, have norm:

P in only I_1: $|Ev_P|^2 = \frac{4}{3}$

P in only I_2: .... $= \frac{4}{3}$

P in both I_1 and I_2 : .... $= \frac{4}{3}(1+1 - \frac{1}{2}\ 2) = \frac{4}{3}$.

The inner product becomes:

P,Q in only I_1 or I_2 $\langle Ev_P, Ev_Q\rangle = \frac{4}{3}$

P in I_1, Q in both I_1 and I_2 .... $= \frac{4}{3}\ (1+0-\frac{1}{2}\ (1+0)) = \frac{2}{3}$

P in I_1 , Q in I_2 but not in both: ...... $= -\frac{1}{2} \cdot \frac{4}{3} = -\frac{2}{3}$

P and Q in both I_1 and I_2 ...... $= \frac{4}{3}\ (1+1-1) = \frac{4}{3}$

Consequently, we have constructed a covariance function with 3 correlation values: 100 %, 50 % and -50 %.

The introduction of one more block gives an even more reasonable function. The third block is supposed to be disjoint with the first, so

$$\bar{I}_3 = (I_3 - \langle\bar{I}_2, I_3\rangle\bar{I}_2)/|(I_3 - \langle\bar{I}_2, I_3\rangle\bar{I}_2)|.$$

Now,

$$\tilde{I}_3 = I_3 - \langle\bar{I}_2, I_3\rangle\bar{I}_2 = I_3 - \frac{2}{\sqrt{3}}\,\frac{1}{2}\,\frac{2}{\sqrt{3}}\,(I_2 - \frac{1}{2}\,I_1)$$

$$= I_3 - \frac{2}{3}\,I_2 + \frac{1}{3}\,I_1$$

$$|\tilde{I}_3|^2 = |I_3|^2 - \frac{2}{3}\,\langle I_3, I_2\rangle - \frac{2}{3}\,\langle I_2, I_3\rangle - \frac{2}{9}\,\langle I_2, I_1\rangle$$

$$+ \frac{4}{9}\,|I_2|^2 - \frac{2}{9}\,\langle I_1, I_2\rangle + \frac{1}{9}\,|I_1|^2 = \frac{2}{3}$$

$$\bar{I}_3 = (\frac{3}{2})^{1/2}\,(I_3 - \frac{2}{3}\,I_2 + \frac{1}{3}\,I_1)\ .$$

Then after some rearrangement

$$K_3(P,Q) = \frac{3}{2}\,(I_1(P)I_1(Q) + 2I_2(P)I_2(Q) + I_3(P)I_3(Q)$$

$$-(I_3(P)I_2(Q) + I_3(Q)I_2(P)) + \frac{1}{2}(I_1(P)I_3(Q) + I_1(Q)I_3(P))$$

$$-(I_1(P)I_2(Q) + I_1(Q)I_2(P))$$

$$|Ev_P|^2 = \frac{3}{2} \text{ everywhere}\ .$$

For the inner product, there are several possibilities:

P in I_1, Q in I_3 $\langle Ev_P, Ev_Q \rangle = \frac{1}{2}$

P in I_1, Q in I_2 and I_3 .... $= -\frac{1}{2}$

P in I_2 and I_1, Q in I_2 and I_3 .... $= 3-1 + \frac{1}{2} - 1 = \frac{3}{2}$

P in I_1 and Q in I_2 and I_1 .... $= \frac{3}{2} - 1 = \frac{1}{2}$

The correlations are then, 100 %, 33 %, -33% and 33 % as a function of distance, as shown in Fig. 7.

References.

Backus,G.E. and J.G.Gilbert: Numerical Applications of a Formalism for Geophysical Inverse Problems. Geophys.J.R.astr.Soc., Vol. 13, pp.247-276, 1967.

Forsberg,R., P.Knudsen and C.C.Tscherning: Description of the GRAVSOFT package. In preparation, 1988.

Hein,G.W., B.Eissfeller, K.Hehl and H.Landau: Processing of Geodetic, Geophysical and satellite data by the new opera 2.3 software system.Presented IUGG XIX General Assembly, Vancouver, Aug. 1987.

Jackson, D.D.: The use of a priori data to resolve non-uniqness in linear inversion. Geophys. J. R. astr. Soc., Vol. 57, pp. 137-157, 1979.

Knudsen, P.: Estimation and Modelling of the Local Empirical Covariance Function using gravity and satellite altimeter data.Bulletin Geodesique, Vol. 61, pp. 145-160, 1987a.

Sanso, F.: Statistical methods in physical geodesy. In: Suenkel, H.: Mathematical and Numerical Techniques in Physical Geodesy. Lecture Notes in Earth Sciences, Vol. 7, pp. 49-155, Springer-Verlag, 1986.

Sanso',F., R.Barzaghi and C.C.Tscherning: Choice of Norm for the Density Distribution of the Earth. Geoph.Jour.Royal Astr. Soc., Vol. 87, pp. 123-141, 1986.

Sanso, F. and C.C.Tscherning: Mixed Collocation: A proposal. Quaterniones Geodasiae, Vol. 3, no. 1, pp. 1-15, 1982.

Tscherning, C.C.: Representation of Covariance Functions Related to the Anomalous Potential of the Earth using Reproducing Kernels. The Danish Geodetic Institute Internal Report No. 3, 1972a.

Tscherning, C.C.: Functional Methods for Gravity Field Approximation. Lecture Notes in Earth Sciences, Vol. 7, pp. 3-47, Mathematical and Numerical Techniques in Physical Geodesy, Springer-Verlag, 1986b.

Tscherning, C.C. and R.H.Rapp: Closed Covariance Expressions for Gravity Anomalies, Geoid Undulations, and Deflections of the Vertical Implied by Anomaly Degree-Variance Models. Reports of the Department of Geodetic Science No. 208, The Ohio State University, Columbus, Ohio, 1974.

Tscherning, C.C. and G.Strykowski: Quasi-harmonic inversion of gravity field data. Presented 5 Int. Seminar Model Optimization in Exploration Geophysics, Berlin, Feb. 4-7, 1987.

Tscherning, C.C. and H. Suenkel: A Method for the Construction of Spheroidal Mass Distributions consistent with the harmonic Part of the Earth's Gravity Potential. Manuscripta Geodaetica, Vol. 6, pp. 131-156, 1981.