High-degree spherical harmonic analysis combining gridded and random distributed data sets

T. Auzinger and W.-D. Schuh[1]

Camera-ready Copy for
Phys. Chem. Earth, Vol. 23, No. 1, pp. 19-23, 1998
Manuscript-No. G12-003

[1]Institute of Theoretical Geodesy, Graz, Austria

Journal: Phys. Chem. Earth, Vol. 23, No. 1, pp. 19-23, 1998
MS No.: G12-003
First author: Auzinger

High-degree spherical harmonic analysis combining gridded and random distributed data sets

T. Auzinger and W.-D. Schuh[1]

[1]Institute of Theoretical Geodesy, Graz, Austria

Received 25 April 1997 – Accepted 1 September 1997

Abstract. A rigorous least squares adjustment of high degree spherical harmonics yields a very large normal equation system. This presentation reports on the numerical problems to solve this system and to obtain accuracy estimates by error propagation. In both cases an optimal reordering strategy makes it possible to decompose the normal equation system in a very efficient way (no fill-in elements). A special adapted Cholesky-factorization and partial inversion algorithm form a well tailored strategy to benefit from this special sparse structure.

1 Introduction

The mathematical representation of earth's gravity potential is conveniently done with normalized spherical harmonics,

$$V(r,\theta,\lambda) = \frac{GM}{r}\left\{1+ \sum_{\ell=1}^{\ell_{max}}\sum_{m=0}^{\ell}\left(\frac{a}{r}\right)^{\ell}\bar{P}_{\ell m}(\cos\theta)\left(\bar{C}_{\ell m}\cos m\lambda+\bar{S}_{\ell m}\sin m\lambda\right)\right\}. \quad (1)$$

GM ... geocentric gravitational constant
a ... semi-major axis
r ... radius vector
θ ... polar distance (colatitude)
λ ... longitude
ℓ ... degree
m ... order
ℓ_{max} ... maximum degree
$\bar{P}_{\ell m}$ ... fully normalized associated Legendre functions of first kind
$\bar{C}_{\ell m}$, $\bar{S}_{\ell m}$ ... harmonic coefficients

All groups of observations can be expressed with the help of the zero, first or second order derivative of the potential. Since the early 1980s a growing number of spherical harmonic models of the earth's gravity field has become available up to very high degrees and orders. Expansions up to degree and order 180 or 360, the latter with more than 130 000 individual coefficients, are widely used for many purposes. All the applications of spherical harmonics can be divided into two tasks:

- determination of the harmonic coefficients $\bar{C}_{\ell m}$ and $\bar{S}_{\ell m}$ using observed quantities of the gravity potential (*'global spherical harmonic analysis'*) and
- computation of the gravity potential and/or derivatives thereof at a special point P (r_P, θ_P, λ_P) with known harmonic coefficients $\bar{C}_{\ell m}$ and $\bar{S}_{\ell m}$ (*'global spherical harmonic synthesis'*).

Due to types and accuracies of the observations, but also due to the point location and distribution of the measurements a more or less high degree expansion can be determined using least squares techniques or analytic approaches.

In spite of new developments in computer design, the standard least squares procedure for the estimation of spherical harmonic coefficients for high degree gravitational fields exceeds present capabilities. Only gridded data can be computed, because of orthogonality relations a sparse normal equation system results. In general, observations can not be performed on a regular grid and especially in connection with orbit dependent measurements no gridded data sets are established. Therefore, the least squares harmonic analysis results in a dense normal equation system. The knowledge about the sparse structure of gridded data can be used for tailored preconditioning strategies. An optimal reordering of the sparse normal equation system brings benefit for gridded data in a direct approach as well as for regular distributed, but not gridded data in an iterative approach. This paper reports on an optimal reordering scheme for high degree spherical harmonic analysis. This optimal reordering strategy brings about the possibility to decompose (e.g. Cholesky-factorization) the normal equations without any fill-in. In addition it is shown, that also the partial inverse can be computed without any additional storage requirement.

Correspondence to: W.-D. Schuh

2 Numbering Schemes

Due to commutativity law of addition it's our free decision how we number the spherical harmonic coefficients $\bar{C}_{\ell m}$ and $\bar{S}_{\ell m}$. Figure (1) illustrates the pattern of all

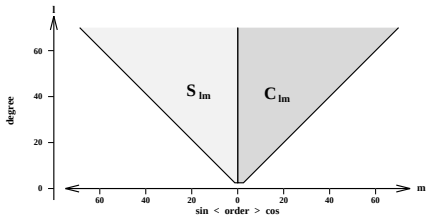

Fig. 1. Triangular scheme with information about the coefficients of the spherical harmonics.

spherical harmonic coefficients. The triangular scheme can be divided into two sub-triangles. The left triangle contains the sine coefficients $\bar{S}_{\ell m}$ while the right triangle contains the cosine coefficients $\bar{C}_{\ell m}$. The degree ℓ increases from bottom to top, and the order m increases from the center to the left (right) for the sine (cosine) coefficients.

Many common numbering schemes are organized by degree, where it is easy to cut-off higher (lower) degrees and orders at ones. Numbering schemes by order organize the normal equations in an optimal way with respect to gridded data sets, because all non-vanishing elements are organized in a block-diagonal structure.

2.1 Kite Numbering Scheme

For optimized combination of dense low degree data and sparse high degree data the following numbering scheme is introduced. First we have to divide the triangular scheme in three parts with respect to the highest degree of dense information (cf. Figure 2).

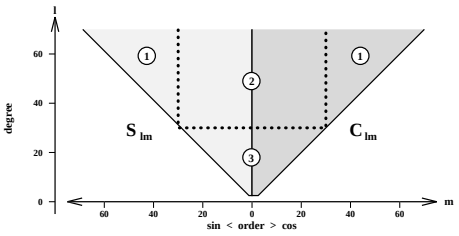

Fig. 2. Kite numbering scheme.

First block #1 is handled. The coefficients are numbered by order. Odd and even coefficients are separated additionally. Cosine and Sine coefficients follow each other for corresponding order. Afterwards block #2 and #3 are handled similarly. In *Schuh (1996b)*, chapter 2 it is shown that no fill-in elements appear during Cholesky's factorization using this structure (cf. Fig. 3).

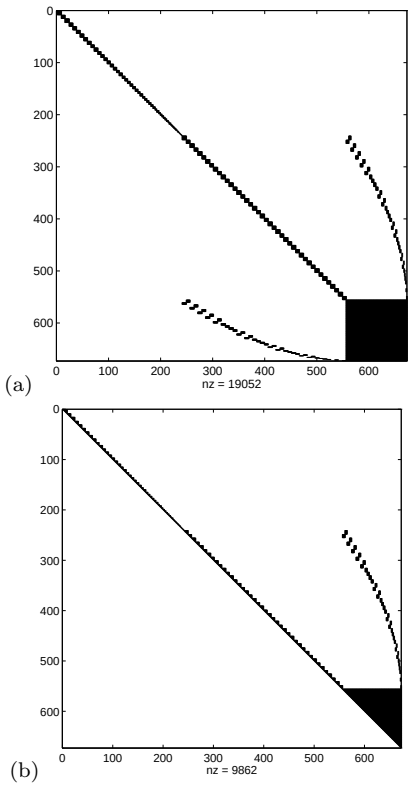

Fig. 3. Sparsity of the (reduced) normal equations employing kite numbering scheme. Block-diagonal structure up to degree 25, full system up to degree 10. (a) normal equations (b) reduced normal equations.

3 Cholesky's Algorithm

The determination of coefficients of earth's gravity field by a global spherical harmonic analysis leads to very large equation system. Since this equation system is usually overdetermined a least squares adjustment is an appropriate way to solve them. The normal equation matrix $\boldsymbol{N}$ related to the system

$$\boldsymbol{N}\boldsymbol{x} = \boldsymbol{n} \tag{2}$$

is due to the least squares model symmetric and positive (semi-) definite. Therefore, Cholesky's method by factorizing the matrix $\boldsymbol{N}$ in an upper triangular matrix $\boldsymbol{R}$ and it's transposed

$$\boldsymbol{N} = \boldsymbol{R}^T\boldsymbol{R} \tag{3}$$

suggests itself. After substituting eq. (3) in eq. (2) the equation system can be split into two triangular systems

$$\underbrace{\boldsymbol{R}^T\underbrace{\boldsymbol{R}\boldsymbol{x}}_{\boldsymbol{z}} = \boldsymbol{n}}_{\boldsymbol{R}^T\boldsymbol{z} = \boldsymbol{n}} \quad . \qquad \Longrightarrow \qquad \boldsymbol{R}\boldsymbol{x} = \boldsymbol{z} \tag{4}$$

The solution is performed in two steps: First the matrix $\boldsymbol{N}$ is factorized to $\boldsymbol{R}$. Within this step, which is

always done in an implicit way, the right hand side $\boldsymbol{n}$ is reduced to the auxiliary vector $\boldsymbol{z}$. This corresponds to the solution of the lower triangular system

$$\boldsymbol{R}^T \boldsymbol{z} = \boldsymbol{n} \tag{5}$$

In a second step the upper triangular system

$$\boldsymbol{R}\, \boldsymbol{x} = \boldsymbol{z} \tag{6}$$

is solved. The first step is performed top down while the second step is done bottom up. Cholesky's algorithm for the solution of a n-dimensional equation system can be summarized in the following way:

$$r_{ii} = \sqrt{n_{ii} - \sum_{k=1}^{i-1} r_{ki}^2} \qquad i = 1, \ldots, n \tag{7}$$

$$r_{ij} = \left(n_{ij} - \sum_{k=1}^{i-1} r_{ki} r_{kj} \right) / r_{ii} \qquad \begin{matrix} i = 1, \ldots, n \\ j = 1, \ldots, i{-}1 \end{matrix} \tag{8}$$

$$z_i = \left(n_i - \sum_{k=1}^{i-1} r_{ki} z_k \right) / r_{ii} \qquad i = 1, \ldots, n \tag{9}$$

$$x_i = \left(z_i - \sum_{k=i+1}^{n} r_{ik} x_k \right) / r_{ii} \qquad i = n, \ldots, 1 \tag{10}$$

The main part of work within the factorization is done in formula (8). A graphical representation (cf. Fig. 4) illustrates this process. The kernel of the computation represents the scalar product (,) of the reduced parts of column i and j. This product is subtracted from the un-reduced element n_{ij} and divided by the already reduced diagonal element r_{ii}.

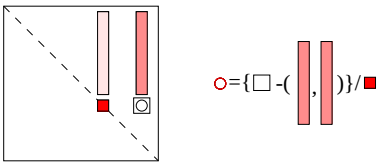

Fig. 4. Cholesky reduction of the element r_{ij} (Eq. (8). In addition to the un-reduced element n_{ij} the column vectors r_{ki}, r_{kj} and the diagonal element r_{ii} of the reduced system are involved.

If we consider gridded data the matrix $\boldsymbol{N}$ is a sparse matrix. In general each factorization step destroys the sparsity of a matrix by changing zero elements to non-zero elements (*'fill-in'*). The apriori knowledge that an element of $\boldsymbol{N}$ keeps zero during factorization frees us from allocating storage requirements and avoids it's computation. It's worth now to focus on the property of a zero element that keeps unchanged during factorization. The following conditions can be stated:

1. the element n_{ij} has to be zero, and
2. the scalar product of the column vectors $r_{ki} r_{kj}$ has to vanish.

The second condition holds if one column vector is zero, or always one of the two elements forming the product is zero as illustrated by figure 5.

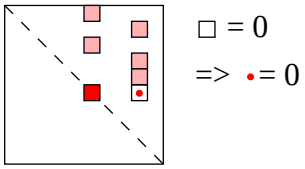

Fig. 5. Cholesky reduction of an element r_{ij} (marked by a square). Since the scalar product of the column vectors vanishes, no fill-in is produced (dot).

Typical examples like block-diagonal, band and profil structured matrices are summarized within figure 6.

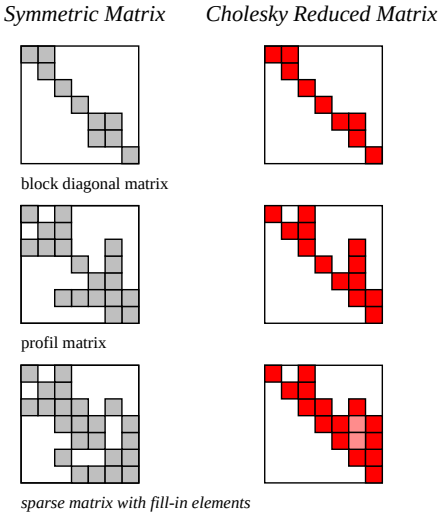

Fig. 6. Typical forms of sparse matrices and their behaviour during a factorization step. Left side: un-reduced form; right side: reduced version.

But we have also in mind that the special structure of figure 7 allows also a factorization without any fill-in element due to vanishing scalar products (cf. Fig. 5).

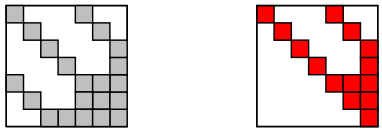

Fig. 7. Special structure without any fill-in elements. Left side: un-reduced form; right side: reduced version.

4 Inverse of the matrix $\boldsymbol{N}$

Usually the computation of the inverse of a matrix needs a huge effort, but is not necessary for the solving an

equation system. However, if least squares adjustment techniques are applied then the inverse of the normal equations matrix holds important statistic information, namely the variances and covariances information of the estimated parameters. From this point of view it is often of interest to determine special parts of the inverse, especially the diagonal or block diagonal part, bearing in mind that the inverse of a sparse matrix is in general a full matrix. This section explains a method to compute a partial inverse, this means, to compute only these elements of the inverse where non-zero elements in the reduced equation system occurs. The base idea of this algorithm was published by *Hanson (1978)*. The utilization of this technique on the kite matrix derived from the harmonic analysis (cf. Fig. 2) allows a very efficient computation of the partial inverse. This computation can be organized *in place* that means that no additional storage requirements occur.

The algorithm can be derived by blocking the matrices $\boldsymbol{N}$ and $\boldsymbol{R}$, respectively.

$$\begin{aligned} \boldsymbol{N} &= \begin{bmatrix} \boldsymbol{N}_{11} & \boldsymbol{N}_{12} \\ (\boldsymbol{N}_{12})^T & \boldsymbol{N}_{22} \end{bmatrix} \qquad (11) \\ \boldsymbol{R} &= \begin{bmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} \\ & \boldsymbol{R}_{22} \end{bmatrix} \\ \boldsymbol{N}^{-1} &= \begin{bmatrix} \boldsymbol{N}_{11}^{(-1)} & \boldsymbol{N}_{12}^{(-1)} \\ \left(\boldsymbol{N}_{12}^{(-1)}\right)^T & \boldsymbol{N}_{22}^{(-1)} \end{bmatrix} . \end{aligned}$$

The blocks $\boldsymbol{N}_{11}^{(-1)}$, $\boldsymbol{N}_{12}^{(-1)}$ and $\boldsymbol{N}_{22}^{(-1)}$ of the inverse matrix $\boldsymbol{N}^{-1}$ are marked by `'(-1)'` in contrast to the inverse `'-1'` of the block itself. These blocks can be computed from the blocks of $\boldsymbol{N}$ in different ways, e.g. for.

$$\boldsymbol{N}_{22}^{(-1)} = \left(\boldsymbol{N}_{22} - \boldsymbol{N}_{12}^T \boldsymbol{N}_{11}^{-1} \boldsymbol{N}_{12}\right)^{-1} \qquad (12)$$

$$\boldsymbol{N}_{12}^{(-1)} = -\boldsymbol{N}_{11}^{-1} \boldsymbol{N}_{12} \boldsymbol{N}_{22}^{(-1)} \qquad (13)$$

$$\boldsymbol{N}_{11}^{(-1)} = \boldsymbol{N}_{11}^{-1} - \boldsymbol{N}_{12}^{(-1)} \boldsymbol{N}_{12}^T \boldsymbol{N}_{11}^{-1} . \qquad (14)$$

This form is chosen, because the Cholesky's reduction in block matrix form shows a very similar behaviour, especially for the $\boldsymbol{R}_{22}$ term

$$\boldsymbol{R}_{22} = \left(\boldsymbol{N}_{22} - \boldsymbol{N}_{12}^T \boldsymbol{N}_{11}^{-1} \boldsymbol{N}_{12}\right)^{C} \qquad (15)$$

$$\boldsymbol{R}_{12} = \left(\boldsymbol{R}_{11}^T\right)^{-1} \boldsymbol{N}_{12} \qquad (16)$$

$$\boldsymbol{R}_{11} = \boldsymbol{N}_{11}^{C} , \qquad (17)$$

where *'C'* means the Cholesky reduced of a matrix. Therefore, it is easy to express (cf. *Schuh (1996a)*, Ch. 2.4.2) the blocks of the inverse $\boldsymbol{N}_{11}^{(-1)}$, $\boldsymbol{N}_{12}^{(-1)}$ (12-14) by the Cholesky reduced blocks $\boldsymbol{R}_{11}$, $\boldsymbol{R}_{12}$, $\boldsymbol{R}_{22}$ (15-17),

$$\boldsymbol{N}_{22}^{(-1)} = \left(\boldsymbol{R}_{22}^T \boldsymbol{R}_{22}\right)^{-1} \qquad (18)$$

$$\boldsymbol{N}_{12}^{(-1)} = -\boldsymbol{R}_{11}^{-1} \boldsymbol{R}_{12} \boldsymbol{N}_{22}^{(-1)} \qquad (19)$$

$$\boldsymbol{N}_{11}^{(-1)} = \left(\boldsymbol{R}_{11}^T \boldsymbol{R}_{11}\right)^{-1} - \boldsymbol{N}_{12}^{(-1)} \boldsymbol{R}_{12}^T \boldsymbol{R}_{11}^{-1} . \qquad (20)$$

This set of equations can be used to elaborate an recursive algorithm. Starting from the element $n_{nn}^{(-1)}$ in the lower right of the matrix in a row-by-row backspace approach the whole inverse matrix can be computed.

Eq. (21)-(23) summarize the scalar form of this algorithm.

$$n_{nn}^{(-1)} = \frac{1}{r_{nn}^2} \qquad (21)$$

$$n_{ij}^{(-1)} = -\frac{1}{r_{ii}} \sum_{k=i+1}^{n} r_{ik} n_{kj}^{(-1)} , \qquad \begin{matrix} j=i+1,\ldots,n \\ i=n,\ldots,1 \end{matrix} \qquad (22)$$

$$n_{ii}^{(-1)} = \frac{1}{r_{ii}^2} - \frac{1}{r_{ii}} \sum_{k=i+1}^{n} r_{ik} n_{ik}^{(-1)} , \; i=n,\ldots,1 \qquad (23)$$

Again the main step of this algorithm the computation of the element $n_{ij}^{(-1)}$ (cf. Eq. (22)) is figured out graphically. Fig. 8 shows that the scalar product of the lower part of the j^{th} column of the inverse with the Cholesky reduced i^{th} row forms the major step. A row-to-row backspace strategy makes sure that all the necessary inverse element in the lower part are computed before.

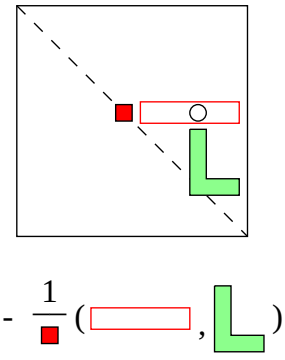

Fig. 8. Graphical representation of eq. (22). Computation of the inverse element $n_{ij}^{(-1)}$ from Cholesky reduced matrix.

It is now a well-known effect, that in general the inversion turns over a sparse system to a dense one. Only a block-diagonal structure is preserved (cf. Fig. 9). But if we are interested only in special elements of the inverse, e.g. the main diagonal elements, or all elements within a defined bandwidth then special algorithms exists to compute these elements in an efficient way. A special algorithm for profile structured matrices was introduced by *Hanson (1978)*. He shows, that all the inverse elements within the profile can be computed in a strict manner without any information of inverse elements outside of the profile. It's up to the fact that the

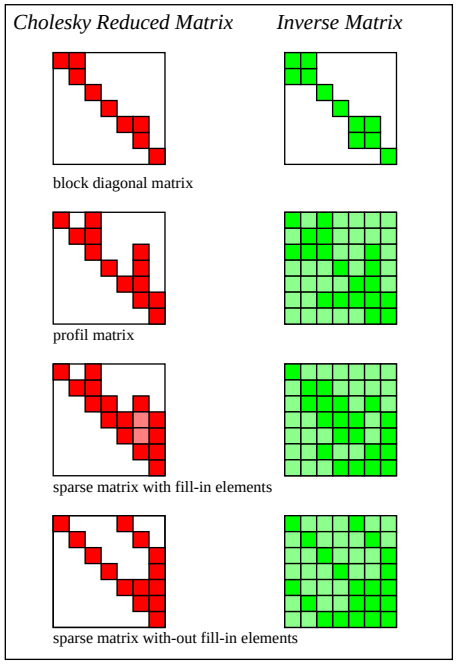

Fig. 9. For basic examples on fill-in elements during Inversion from Cholesky reduced matrix. The darker elements in the Inverses can be computed separately.

non-computed elements within each column of the inverse corresponds exactly with the zero-elements within the Cholesky reduced row. Therefore, the scalar product (Eq. (22), see also Fig. 8) for all elements within the profile is not influenced by inverse element, that are positioned outside of the profile. This argument holds also for our special *'kite'* structure (cf. Fig. 2). This means we can compute the inverse elements for all non-zero elements, denoted as (*'partial inverse'*), in a very efficient *'in place'* strategy. Fig. 10 illustrates the fact that all

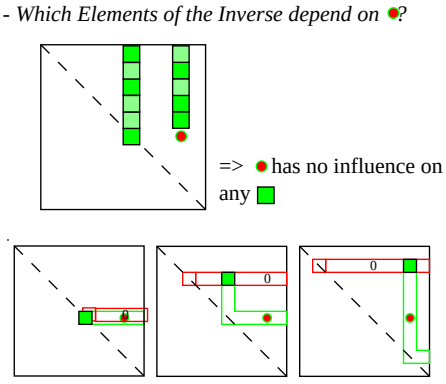

Fig. 10. Computation of the partial inverse from the Cholesky reduced matrix. The inverse element $n_{ij}^{(-1)}$ (dot) influences the inverse values in the i^{th} and j^{th} columns. Since the factorization produces no fill-in elements (cf. 5) the influence of the inverse element $n_{ij}^{(-1)}$ on all lighter inverse elements runs into zero products.

zero-elements after Cholesky reduction (due to vanishing inner products, cf. Fig. 5) do not influence the computation of the partial inverse of all non-zero elements, since the always run into zero products.

5 Concluding Remarks

In connection with tailored numerical strategies the kite numbering scheme turns out as a very powerful strategy to overcome the numerical problems of high degree spherical harmonic analysis. This numbering scheme, especially for the combination of dense low degree fields and block-structured high degree fields produce no fill-in during the factorization step. But also the partial inverse can be computed in an very efficient way. Therefore, this strategy can be applied very efficient to the following scenarios:

1. in case of a combined data set (ungridded low frequency data and gridded high frequency data) a very efficient direct solution and error estimation can be determined and
2. in the case of ungridded low frequency data and irregular distributed high frequency data, which densely cover the area between two parallels, this strategy can be used as optimal preconditioning technique.

The fast iterative solution mainly bases on the possibility to find a representative matrix, which can be easily computed, quickly solved and on the other side represents as close as possible all informations of the whole system. If the block structured high degree model combined with a dense low degree model is used as a mask for the representative matrix both facts are fulfilled, easiest computation and a close relation between the behaviour of the rigorous dense system and the sparse preconditioner.

Numerical simulations demonstrate that these iterative techniques are well tailored and ideally suited for the least squares solution of harmonic coefficients of high degree models based on satellite-to-satellite tracking data and satellite gravity gradiometry data.

References

HANSON, Robert H. (1978): A Posteriori Error Propagation, Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, Arlington, Virginia.

SCHUH, Wolf-Dieter (1996a): Numerische Verfahren zur geodätischen Optimierung. Institut für Theoretische Geodäsie, Abteilung für Mathematische Geodäsie und Geoinformatik, Technische Universität Graz, Graz.

SCHUH, Wolf-Dieter (1996b): Tailored Numerical Solution Strategies for the Global Determination of the Earth's Gravity Field. Mitteilungen der geodätischen Institute der TU-Graz, Folge 81, Graz.