

# Helmert blocking algebra in ODAS

Alex Bombrun

GAIA-C3-TN-ARI-APB-002-01

December 11, 2008

## Abstract

This technical note describes the mathematical process used to solve a least square problem in space astrometry that presents an Helmert blocking structure. We focus particularly on its implementation in ODAS. We assume that the reader is familiar with the space astrometric least square problem. This document is based on [1] and [2].

## 1 Mathematical description

### 1.1 Astrometric observations

In this document we refer to an observation with the term transit time  $t$ . For ODAS an observation is an observed field angle. It can be the along scan  $\eta$  or the across scan  $\zeta$  field angle. Note that each elementary contains along-scan information, but for some transits, across-scan information may be missing. It is important that the code can handle both cases.

Each observed transit time  $t_{i,j,k}$  is associated with a specific object (e.g., a primary source  $i$  with astrometric parameters  $s_i$ , for ODAS this is the source position  $(\epsilon, R)$  in the reference great circle frame, and with certain attitude ( $a_j$ ) and calibration ( $c_k$ ) parameters. A non-linear observation equation is associated with each observation:

$$t_{i,j,k} - f(s_i, a_j, c_k) = e_{i,j,k}, \quad (1)$$

$e_{i,j,k}$  is the error due to imprecision in the model  $f$ , measurement errors, and so on. In particular one priori estimation of the error  $\sigma_{i,j,k}$  is given for each observation. These estimations will be used to balance the observation equations. The error should be small if our knowledge of the source, attitude and calibration parameters is good. Note that  $f(s, a, c)$  is a highly non-linear but smooth function. Hence we will use a linear approximation around some initial reference values  $\bar{s}_i, \bar{a}_j, \bar{c}_k$ . Let  $h_{s_i}, h_{a_j}, h_{c_k}$  be the displacements around these reference values. Once these values will have been computed they will be called updates, and the updated value will be  $s_i = x\bar{s}_i + h_{s_i}$ ,  $a_j = \bar{a}_j + h_{a_j}$ ,  $c_k = \bar{c}_k + h_{c_k}$ .

Considering all observations, we get an over-determined system of linear equations in  $h$ , namely the observation equations

$$\forall t_{i,j,k} : \quad S_{i,j,k} h_{s_i} + A_{i,j,k} h_{a_j} + C_{i,j,k} h_{c_k} = r_{i,j,k} (+e_{i,j,k}) \quad (2)$$

with the residuals  $r_{i,j,k} = t_{i,j,k} - f(\bar{s}_i, \bar{a}_j, \bar{c}_k)$  and the Jacobian matrices  $S_{i,\tau} = (\partial f / \partial s_i)(\bar{s}_i, \bar{a}_j, \bar{c}_k)$ ,  $A_{i,\tau} = (\partial f / \partial a_j)(\bar{s}_i, \bar{a}_j, \bar{c}_k)$ ,  $C_{i,\tau} = (\partial f / \partial c_k)(\bar{s}_i, \bar{a}_j, \bar{c}_k)$ . A weight matrix  $W_i$  associated with the source  $i$ , it is a diagonal matrix whose length is set by the number of observations available for the source  $i$ . Each element on the diagonal is a positive real number equal to  $\sigma_{i,j,k}^2 / \sigma_0^2$ , where  $\sigma_0$  is a positive global (i.e. valid for all sources) parameter.

Note that the astrometric observation equation for Gaia is set using two functions  $f_1$  and  $f_2$ . More precisely  $f_1(t_{i,j,k}, c_k) - f_2(s_i, a_j) = e_{i,j,k}$ , hence a factor minus appears in front of  $C_{i,j,k} = \partial f_1 / \partial c_k$  when setting the astrometric design equation for Gaia,  $S_{i,j,k} h_{s_i} + A_{i,j,k} h_{a_j} - C_{i,j,k} h_{c_k} = r_{i,j,k}$ .

### 1.2 Triangular structure

An important point is that there are many more source parameters than attitude and calibration parameters. From Eq. (2) we see that the observation equations for different sources involve disjoint source parameter vectors  $s_i$ ,

while the attitude and calibration vectors  $a_j, c_k$  may overlap. Sorting all the observation equations (2) by the source index  $i$  and collecting them in one matrix we get a non-square block angular matrix:

$$\begin{bmatrix} S_1 & 0 & \dots & 0 & O_1 \\ 0 & S_2 & \dots & 0 & O_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & S_n & O_n \end{bmatrix} \begin{bmatrix} h_{s_1} \\ h_{s_2} \\ \vdots \\ h_{s_n} \\ h_o \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \Leftrightarrow Kh = r \quad (3)$$

with  $n$  being the number of primary sources, the matrices and vector  $S_i, O_i, r_i$  being the concatenation of, respectively, all  $S_{i,j,k}, O_{i,j,k} = [A_{i,j,k}, C_{i,j,k}], r_{i,j,k}$ , and  $h_o$  being the vector of all attitude and calibration displacement unknowns  $h_o = [h_a, h_c]^T$ .

### 1.3 The reduced normal matrix

The system (3) is over-determined: there are more equations than unknowns. Due to measurement errors, there does not exist a solution that simultaneously satisfies all the equations. However, the problem becomes mathematically well posed when we try to minimise the norm of the post-fit residual vector,  $\|r - Kh\|$ . This is the ordinary least-squares problem, which is classically solved by forming the normal equations

$$K^T W K h = K^T W r \quad (4)$$

or

$$\begin{bmatrix} S_1^T W_1 S_1 & 0 & \dots & 0 & S_1^T W_1 O_1 \\ 0 & S_2^T W_2 S_2 & \dots & 0 & S_2^T W_2 O_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & S_n^T W_n S_n & S_n^T W_n O_n \\ O_1^T W_1 S_1 & O_2^T W_2 S_2 & \dots & O_n^T W_n S_n & \sum_i O_i^T W_i O_i \end{bmatrix} \begin{bmatrix} h_{s_1} \\ h_{s_2} \\ \vdots \\ h_{s_n} \\ h_o \end{bmatrix} = \begin{bmatrix} S_1^T W_1 r_1 \\ S_2^T W_2 r_2 \\ \vdots \\ S_n^T W_n r_n \\ \sum_i O_i^T W_i r_i \end{bmatrix} \quad (5)$$

A standard way to handle normal equations with the block-diagonal-bordered structure of Eq. (5) is to successively eliminate the unknowns along the block-diagonal (in our case the source parameters), leaving us with a reduced normal equations system for the remaining unknowns (in our case the attitude and calibration parameters). A straightforward computation shows that the solution of Eq. (5) can be accomplished by first solving the reduced normal equations for the attitude parameters,

$$\left[ \sum_{i=1}^n \left( O_i^T W_i O_i - O_i^T W_i S_i (S_i^T W_i S_i)^{-1} S_i^T W_i O_i \right) \right] h_o = \sum_{i=1}^n \left( O_i^T - O_i^T W_i S_i (S_i^T W_i S_i)^{-1} S_i^T \right) W_i r_i \quad (6)$$

and then forwarding the solution  $h_o$  to solve all the source equations

$$h_{s_i} = (S_i^T W_i S_i)^{-1} S_i^T W_i (r_i - O_i h_o). \quad (7)$$

The reduced normal equation (6) will be denoted in a compact formulation by  $M h_o = \mathbf{b}$ , with  $M$  being the reduced normal matrix and  $\mathbf{b}$  the right hand side of the reduced normal equation.

The reduced normal equation is obtained by a sum over all sources. Hence the process of fill-in the reduced normal equation is the same for each source. We have first to compute  $S_i, O_i, W_i$  and  $r_i$ , i.e. building the design equation of the source being processed. Then we have to compute the contribution of the source  $i$  to the reduced normal matrix,

$$\left( O_i^T W_i O_i - O_i^T W_i S_i (S_i^T W_i S_i)^{-1} S_i^T W_i O_i \right), \quad (8)$$

and to the right hand side of the reduced normal equations

$$\left( O_i^T - O_i^T W_i S_i (S_i^T W_i S_i)^{-1} S_i^T \right) W_i r_i. \quad (9)$$

In order to do that efficiently, we have to use the sparseness structure of the design matrix  $O_i = [A_i, C_i]$ .

## 1.4 Constraints

Beside to the design equation the final problem contains linear constraints relative to some of the parameters  $h_o$ . A constraint is simply an other set of linear equations that the parameters should fullfil,

$$Hh_o = r_H. \quad (10)$$

There is two methods to manage linear constraints: introduce fictitious observations or use Lagrange multipliers. Using the fictitious observations you should decide of the weight assigned to the constraints, whereas using the Lagrange multiplier you compute the best fit for the weight. The two methods should be treated differently. In the context of Helmert blocking algebra, if we assume that there is no constraint on the source unknowns, both constraints can be added after the fill-in of the reduced normal matrix.

### 1.4.1 fictitious observations

Let  $\sigma > 0$ , be a weight factor for the constrains. The fictitious observation method consists in adding new observation equation to the design matrix powdered by the weight  $\sigma$ . Since we assume no constraint on the source parameters but only on some of the parameters  $h_o$ , the fictitious observations appears as new terms that have to be added to the reduced normal matrix. Instead of  $\mathbf{M}h_o = \mathbf{b}$ , consider

$$[\mathbf{M} + H^T \sigma^{-1} H] h_o = \mathbf{b} + H^T \sigma^{-1} r_H. \quad (11)$$

### 1.4.2 Lagrange multipliers

Whereas there is no new unknowns using fictitious observations, the number of unknowns is extended according to the number of constraints, i.e. the number of lines in  $H$ , when the constraints are treated using Lagrange multipliers. Again since we assume that only some of the parameters  $h_o$  are constraint, the Lagrange multipliers  $k$  can be added after the computation of the reduced normal matrix. This operation consist in adding new lines and columns to the reduced normal equation. Instead of  $\mathbf{M}h_o = \mathbf{b}$ , consider

$$\begin{bmatrix} \mathbf{M} & H \\ H^T & 0 \end{bmatrix} \begin{bmatrix} h_o \\ k \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ H^T r_H \end{bmatrix}. \quad (12)$$

## 1.5 Statistics

The least squares method is famous due to the Gauss-Markoff theorem. This section summerizes the classical statistical interpretation of the least squares method. The expected value of a random vector  $x$  is denoted by  $E[x]$  and its covariance matrix by  $\nu(x)$ .

Assume that given  $M$  and  $\bar{y}$ , there exists  $\bar{x}$  such that  $\bar{y} = M\bar{x}$ . Assume that the observations are a vector of random variables,  $y = \bar{y} + e$ , with  $e$  the error vector of random variables. The vector  $\bar{x}$  of parameters is related to the vector  $y$  of observations by an over-determinate linear system  $M\bar{x} = y$ . If the expected value of the error is zero,  $E[e] = 0$ , if the covariance matrix of the error,  $\nu(e) = W^{-1}$  with  $W$  positive definite, if the design matrix  $M$  has rank  $n$ , the dimension of the parameter, then the best unbiased linear estimate for the parameter  $\bar{x}$  is the solution,  $\hat{x} = (M^T W M)^{-1} M^T W y$ , of the weighted least squares problem,  $\min_x (Mx - y)^T W (Mx - y)$ . A random vector  $\hat{x}$  function of the observation  $y$  is an unbiased estimate of the parameter  $\bar{x}$  if  $E[\hat{x}] = \bar{x}$ .

The variance matrix  $\nu(\hat{x})$  of the estimator  $\hat{x}$  is linked to the covariance matrix of the observations  $\nu(y) = W^{-1}$  and the design matrix  $M$

$$\nu(\hat{x}) = (M^T W M)^{-1}. \quad (13)$$

The variance of the residual  $\hat{r} = y - M\hat{x}$  is linked to the covariance of the observations  $\nu(e) = W^{-1}$  and the design matrix  $M$

$$\nu(\hat{r}) = (I - H)W^{-1}(I - H), \quad (14)$$

with the hat matrix  $H = M(M^T W M)^{-1} M^T$ .

## 2 Toward an implementation

This section details a possible implementation of the computation presented above. This implementation is simply a way to process the algebra using the sparseness structure based on the idea of storing only the non-zero columns. It should be possible to use some sparse matrix libraries instead. But, since we have none knowledge about their efficiency, it is rather preferable not to rely on them.

### 2.1 Notation

Be aware that in this section we remove the reference to the source  $i$  on the quantities  $r$ ,  $O$ ,  $S$  and  $W$ . Any indices on such an object refer now to rows and/or columns.  $r_i$  is the component  $i$  of the vector  $r$ .  $O_{ij}$  is the component on the row  $i$  at the column  $j$  of the matrix  $O$ .

### 2.2 Design equation of a source $i$

#### 2.2.1 design matrix

The design matrix  $[S, O]$  of a sources  $i$  is a sparse matrix. In order to have an efficient implementation it is important to exploit this property. Since the computation of the reduced normal matrix implies manipulation of column vectors without referring explicitly to the rows orders a good choice is to store only the non-zero columns. As we have seen previously the design matrix of a source  $i$  is the concatenation of  $S$  and  $O$ . The matrix  $S$  has  $n_s$  columns,  $n_s$  being the number of source unknowns per source for ODAS its 2. These columns are assume to be non zeros whereas the matrix  $O$  has a more complex structure. In order to construct the design matrix it is sufficient to know the indices and the values of the non-zero columns.

In this section we make the following choice for the indices of the non-zero columns

$$indices = [-n_s, \dots, -1, \dots, index_j, \dots]. \quad (15)$$

We choose negative index for columns corresponding to the matrix  $S_i$  and positive index for matrix  $O_i$ . Moreover we assume that the indices are ordered from the smaller to the larger, i.e. if  $j_1 < j_2$  then  $index_{j_1} < index_{j_2}$ . This choice simplify the fill-in of the reduced normal equation, see the step 3 in Section 2.3 Note that the positive indices  $index_j$  referred to the order chosen for the other unknowns  $h_o$ . If  $n_o$  denotes the number of other unknowns, i.e. the length of the vector  $h_o$ , then  $index_j < n_o$ .

This vector of indices is in one to one correspondence with the non-zero columns of the design matrix  $[S, O]$ :

$$columns = [c_0, \dots, c_{n_s-1}, \dots, c_j, \dots], \quad (16)$$

through the map  $j \rightarrow (index_j, c_j)$  from  $\{0, \dots, n_{columns}\}$  to  $\mathbb{Z} \times \mathbb{R}^{n_{rows}}$ . Each vector  $c_j$ ,  $0 \leq j < n_{columns}$ , has a length  $n_{rows}$  with  $n_{columns}$  the number of non-zero columns in the design matrix of the source  $i$ , and  $n_{rows}$  the number of observation equations for the source  $i$ .

It is important to understand that the indices of the non-zero columns correspond exactly to the unknowns that contribute to the observation equation of the source  $i$ . But at this level of computation we should not care about what they are!

#### 2.2.2 right hand side

The right hand side  $r$  of the design equation is a vector of length  $n_{rows}$  that contains the observed minus computed values.

#### 2.2.3 weight matrix

The weight matrix  $W$  is a vector  $w$  of length  $n_{rows}$  that contains the prior errors made on the observations. i.e. we assume uncorrelated observations with variance  $\sigma_i$ . Moreover we assume that the weight matrix has been normalized, i.e.  $\max w_i = 1$ .

### 2.3 Computation of the contribution of a source $i$ to the reduced normal equation

Once the design equation of the source  $i$  as been set up, the process of computing the contribution of this source to the reduced normal matrix can be done in 3 steps using 2 main variables a matrix  $M$  and a vector  $b$ :

1. setup,

$$M = \begin{bmatrix} S^T W S & S^T W O \\ O^T W S & O^T W O \end{bmatrix}$$

$$b = \begin{bmatrix} S^T W r \\ O^T W r \end{bmatrix}$$

2. reduce.

$$M = \begin{bmatrix} * & * \\ * & (O^T W O - O^T W S (S^T W S)^{-1} S^T W O) \end{bmatrix}$$

$$b = \begin{bmatrix} * \\ (O^T - O^T W S (S^T W S)^{-1} S^T) W r \end{bmatrix}$$

3. fill-in the reduced normal equation composed of the reduced normal matrix  $M$  and the reduced normal right hand side  $b$ .

The vector  $b$  is of length  $n_{columns}$ , the number of non-zero columns in the design matrix of the source  $i$ .

The symmetric matrix  $M$  with  $n_{columns} \times n_{columns}$  elements can also be stored within a vector of length  $n_{columns}(n_{columns} + 1)/2$ . The element  $M_{ij}$  of the upper triangular part of the matrix  $M$  on the row  $i$ ,  $0 \leq i < n_{columns}$ , at the column  $j$ ,  $i \leq j < n_{columns}$  being at the position  $i + j(j + 1)/2$  on the vector,  $0 \leq i + j(j + 1)/2 < n_{columns}(n_{columns} + 1)/2$ .

#### 2.3.1 setup

```

1:  $M \leftarrow 0$  {initialization}
2:  $b \leftarrow 0$ 
3: for  $i = 0; i < n_{columns}; i_{++}$  do
4:    $x \leftarrow c_i$ 
5:   for  $k = 0; k < n_{rows}; k_{++}$  do
6:      $M_{ij} \leftarrow M_{ij} + x_k w_k x_k$ 
7:      $b_i \leftarrow b_i + x_k w_k r_k$ 
8:   end for
9:   for  $j = i + 1; j < n_{columns}; j_{++}$  do
10:     $y \leftarrow c_j$ 
11:    for  $k = 0; k < n_{rows}; k_{++}$  do
12:       $M_{ij} \leftarrow M_{ij} + x_k w_k y_k$ 
13:    end for
14:  end for
15: end for

```

#### 2.3.2 reduce

**Require:**  $n_{columns} \geq n_s$

```

1: for  $i = 0; i < n_s; i_{++}$  do
2:   if  $M_{ii} \neq 0$  then
3:     for  $j = i + 1; j < n_{columns}; j_{++}$  do
4:       for  $k = i + 1; k \leq j; k_{++}$  do
5:          $M_{kj} \leftarrow M_{kj} - M_{ik} M_{ij} / M_{ii}$ 
6:       end for
7:        $b_j \leftarrow b_j - M_{ij} b_i / M_{ii}$ 
8:     end for
9:   end if
10: end for

```

### 2.3.3 fill-in

```

1: for  $i = n_s; i < n_{columns}, i_{++}$  do
2:    $I \leftarrow indices(i)$  {depend on the choice made at (15)}
3:    $\mathbf{b}_I \leftarrow \mathbf{b}_I + b_i$ 
4:   for  $j = i; j < n_{columns}, j_{++}$  do
5:      $J \leftarrow indices(j)$  {depend on the choice made at (15)}
6:      $\mathbf{M}_{IJ} \leftarrow \mathbf{M}_{IJ} + M_{ij}$ 
7:   end for
8: end for

```

## 2.4 Add the constraints

The constraints are specific to the modelization of the problem. As such they should be defined at the same time than the parameters  $h_o$ . As pointed out in Section 1.4 the software should be flexible enough to treat some constraints with the fictitious observation method (11) and some with the Lagrange multipliers method (12). Note that whereas the fictitious observation method does not change the structure of the reduced normal equation,  $\mathbf{M}x = \mathbf{b}$ , the Lagrange multiplier method extends the number of unknowns. Hence, the size of the reduced normal matrix should have been properly defined to handle the constraints.

## 2.5 Solve the reduced normal equation

Any linear solver that returns a solution  $\hat{h}_o$  of the linear equation  $\mathbf{M}x = \mathbf{b}$ . Note that the covariance matrix of the estimator  $\hat{h}_o$  is simply

$$\nu_{\hat{h}_o} = \mathbf{M}^{-1}. \quad (17)$$

## 2.6 Back substitution

Once the solution of the reduced normal matrix as been computed, the back substitution process can start, i.e. for each source compute the result  $\hat{h}_s$  of the matrix product (7). Moreover we are also interested in other quantities that reflect the quality of the updates. These computations are important for the treatment of outliers.

For each sources, we propose the following values:

- Source parameter estimator

$$\hat{h}_s = (S^T W S)^{-1} S^T W (r - O h_o) \quad (18)$$

- Covariance of the source parameter estimator

$$\nu_{\hat{h}_s} = ((S^T W S)^{-1} S^T W) [O \mathbf{M}^{-1} O^T] ((S^T W S)^{-1} S^T W)^T \quad (19)$$

- Residuals estimator

$$\hat{r} = r - O \hat{h}_o - S \hat{h}_s \quad (20)$$

- Covariance of the residuals estimator

$$\nu_{\hat{r}} = W^{-1} - O \mathbf{M}^{-1} O^T - S (S^T W S)^{-1} S^T \quad (21)$$

All these computation can be performed using classical matrix vector operations.

## 3 Remarks

- This document has been designed to support the ODAS development. In a close future it will be included in the ODAS documentation.
- This document does not consider the back-rotation since this process refers to the particular nature of the source.

- As well, we have not discuss the data model (or extended data model) and the constraints used to computed the one day astronomical solution because it is important that the code depends the least possible of such a choice. Our choice has been to separate the Helmert algebraic operations from the specific choice of a data model such that the code can be reused for other computation such that a ring to sphere.

## References

- [1] H.H. Bernstein, U. Bastian, and S. Hirte. First look - description of the ring solution. November 2005.
- [2] Å. Björck. *Numerical Methods for Least Squares Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1996.