Normal Places in Least-Squares problems, and their relation to the Ring Solution and Ring-to-Sphere Solution

prepared by: L. Lindegren
approved by:
reference: GAIA-C3-TN-LU-LL-076-01
issue: 1
revision: 0
date: 2008-10-15
status: Issued

Abstract

This note is an attempt to define precisely the meaning of 'Normal Place' in the context of linear least-squares problems. Under certain conditions, it is found that the use of normal places allows an exact decomposition of the problem. Its relation to the Ring Solution and the Ring-to-Sphere Solution is briefly discussed.

Document History

Issue	Revision	Date	Author	Comment
1	0	2008-10-15	LL	First issue, no significant changes from draft
D	0	2008-07-01	LL	First draft

1 Introduction

The use of 'normal places' (*Normalorte* in German) is ubiquitous in older astronomical literature, e.g., for least-squares determination of the orbits of asteroids, comets and visual binaries. The basic idea is that multiple observations that have been made within a limited time interval (say a single day, month or even year, depending on the period of the orbit) can sometimes be grouped together, and subsequently treated as a single observation with the combined weight of the individual observations. When considering a single coordinate (e.g., declination δ) the method amounts to computing a weighted mean of a group of residuals $(O-C)_\delta$ with respect to some reference orbit, and then treating this mean residual as an observation referring to the weighted mean time of observation within the group (e.g., von Oppolzer, 1880, p. 371). Although the practice of using normal places is probably at least as old as the least-squares method itself, it is seldom explicitly described in textbooks and its theoretical foundation is even more rarely discussed.

Using normal places can be seen as a kind of pre-processing or compression of the data, and as such could greatly reduce the amount of computation needed to process long series of observations. Understandably, the method was therefore popular before the advent of electronic computers. Nowadays it has largely fallen out of use since there is no longer any justification for trying to reduce the number of data points prior to the least-squares adjustment, especially since that procedure is likely to involve some approximations.

A somewhat related concept is the use of one-dimensional 'star abscissae' in the Hipparcos data processing (ESA, 1997, Vol. 3, Ch. 9), and the two-dimensional positions obtained in the so-called Ring Solution (Bernstein et al., GAIA-ARI-BST-001-5; Hirte et al., SH-003) as part of the Gaia First-Look Processing. Thus the concept is by no means alien to the modern thinking and a variant of it could perhaps be useful as a practical way to decompose the full astrometric solution for Gaia, e.g., in terms of the Ring and Ring-to-Sphere solutions. It is therefore interesting to explore the concept a bit further.

In this note I attempt to define the concept of normal places strictly in the context of linear least-squares estimation (Sect. 2). It is shown (Sect. 3) that the procedure, under certain circumstances, is exact in the sense that it produces the same least-squares estimate as a direct processing of the individual observations. An important modification (and generalization) is proposed in Sect. 4. Finally, in Sect. 5 I briefly discuss its application to the more common situation when the decomposition is not exact, with an example from Gaia and a possible application to the Ring and Ring-to-Sphere solutions.

2 Mathematic formulation

We consider the linear least-squares problem of estimating the n-dimensional parameter vector $\boldsymbol{x}$ based on a set of $m > n$ uncorrelated measurements $\boldsymbol{z}$. The observation equations are therefore

$$\boldsymbol{A}\boldsymbol{x} \cong \boldsymbol{z} \tag{1}$$

where $\boldsymbol{A}$ is a known coefficient matrix of dimension $m \times n$ and rank n. We will use index $i = 1 \ldots n$ for the parameters and $k = 1 \ldots m$ for the observations; thus the observation equations in (1) can also be written

$$\forall k : \qquad \sum_i A_{ki} x_i \cong z_k \tag{2}$$

Here and in the following we use the convention that a subscripted summation symbol implies a summation over the full range of that subscript.

Without loss of generality we may assume that all the observations have the same weight – otherwise, just multiply each equation in (2) by the square root of the weight of the corresponding observation (or divide by its standard deviation). The direct least-squares estimate of the parameter vector is

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z} \tag{3}$$

We shall now assume that the observations k are divided into subsets K_j, $j = 1 \ldots J$, subsequently called groups, such that each k belongs to exactly one group. The original problem (1) can now be divided up into the J partial problems

$$\forall j : \qquad \boldsymbol{A}_j\boldsymbol{x} \cong \boldsymbol{z}_j \tag{4}$$

where $\boldsymbol{A}_j$ is the sub-matrix of coefficients obtained by deleting all rows in $\boldsymbol{A}$ except those belonging to group j, and similarly $\boldsymbol{z}_j$ is the sub-vector of measurements belonging to group j. We use m_j to denote the cardinality of K_j, i.e., the number of equations in (4) for each j. Equation (3) can be written

$$\hat{\boldsymbol{x}} = \left(\sum_j \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{A}_j\right)^{-1} \sum_j \boldsymbol{A}_j^{\mathrm{T}}\boldsymbol{z}_j \tag{5}$$

For each group we now introduce a normal place, $\boldsymbol{y}_j$. Each normal place is a vector of dimension p_j, where $1 \le p_j \le n$. However, for simplicity we assume that the normal places all have the same dimension p. Implicit in the concept of normal places is that $p < n$ (although this is not strictly necessary).

Using normal places, the least-squares adjustment proceeds in two steps: first, consider separately each group of observations K_j and make a least-squares estimation of its normal place $\boldsymbol{y}_j$. In the second step, all the normal places are combined in a least-squares estimation of $\boldsymbol{x}$.

3 Exact decomposition

Under certain conditions, the above procedure using normal places may produce exactly the same estimate of $\boldsymbol{x}$ as the direct solution (3). To see this, suppose that each sub-matrix $\boldsymbol{A}_j$ (of dimension $m_j \times n$) can be written

$$\boldsymbol{A}_j = \boldsymbol{B}_j \boldsymbol{C}_j \tag{6}$$

where $\boldsymbol{B}_j$ is a matrix of dimension $m_j \times p$ and $\boldsymbol{C}_j$ is a matrix of dimension $p \times n$. It is readily seen that this decomposition is possible if and only if $\mathrm{rank}(\boldsymbol{A}_j) \leq p$.

Then consider in the first step the J separate least squares problems

$$\boldsymbol{B}_j \boldsymbol{y}_j \cong \boldsymbol{z}_j \tag{7}$$

Assuming that $\boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j$ is non-singular, the solutions are

$$\hat{\boldsymbol{y}}_j = \left(\boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j\right)^{-1} \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{z}_j \tag{8}$$

The observations in (7) are uncorrelated and of unit weight according to our assumptions, and therefore the covariance of $\hat{\boldsymbol{y}}_j$ is given by $\boldsymbol{V}_j = (\boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j)^{-1}$. Using the Cholesky factorization of the normal matrix,

$$\boldsymbol{L}_j \boldsymbol{L}_j^{\mathrm{T}} = \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j \tag{9}$$

where $\boldsymbol{L}_j$ is a left-triangular matrix, we have $\boldsymbol{V}_j = \boldsymbol{L}_j^{-\mathrm{T}} \boldsymbol{L}_j^{-1}$.

Let us now in the second step consider the 'observation equations'

$$\forall j: \qquad \boldsymbol{C}_j \boldsymbol{x} \cong \hat{\boldsymbol{y}}_j \tag{10}$$

The normal places are mutually uncorrelated (since they are formed from disjoint subsets of the observations), but they have in general covariances $\boldsymbol{V}_j$ that are different from the identity matrix. The observation equations (10) therefore need to be weight-normalized before they are combined in a least-squares solution. Consider the transformed equations obtained by pre-multiplying each equation in (10) with the transposed Cholesky factor:

$$\forall j: \qquad \boldsymbol{L}_j^{\mathrm{T}} \boldsymbol{C}_j \boldsymbol{x} \cong \boldsymbol{L}_j^{\mathrm{T}} \hat{\boldsymbol{y}}_j \tag{11}$$

The covariance of the right-hand side is seen to be

$$\mathrm{Cov}\left[\boldsymbol{L}_j^{\mathrm{T}} \hat{\boldsymbol{y}}_j\right] = \boldsymbol{L}_j^{\mathrm{T}} \mathrm{Cov}\left[\hat{\boldsymbol{y}}_j\right] \boldsymbol{L}_j = \boldsymbol{L}_j^{\mathrm{T}} \boldsymbol{V}_j \boldsymbol{L}_j = \boldsymbol{L}_j^{\mathrm{T}} \boldsymbol{L}_j^{-\mathrm{T}} \boldsymbol{L}_j^{-1} \boldsymbol{L}_j = \boldsymbol{I} \tag{12}$$

Thus we can treat (11) as a set of pJ uncorrelated, unit weight observations. Forming the normal equations for $\boldsymbol{x}$ in the usual manner we find

$$\left(\textstyle\sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{L}_j \boldsymbol{L}_j^{\mathrm{T}} \boldsymbol{C}_j\right) \boldsymbol{x} = \textstyle\sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{L}_j \boldsymbol{L}_j^{\mathrm{T}} \hat{\boldsymbol{y}}_j \tag{13}$$

Inserting (9) and (8) we have

$$\begin{aligned}\left(\sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j \boldsymbol{C}_j\right) \boldsymbol{x} &= \sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j \left(\boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j\right)^{-1} \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{z}_j \\ &= \sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{z}_j \end{aligned} \tag{14}$$

Since $\boldsymbol{B}_j \boldsymbol{C}_j = \boldsymbol{A}_j$ we find

$$\left(\sum_j \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{A}_j\right) \boldsymbol{x} = \sum_j \boldsymbol{A}_j^{\mathrm{T}} \boldsymbol{z}_j \tag{15}$$

which is the same as the direct solution (5).

Thus we have shown that the use of normal places gives identically the same estimate (and covariance of the estimate) as the direct solution provided that:

1. the partial observation equations matrices $\boldsymbol{A}_j$ are factored as in (6)
2. the full covariance matrices of the resulting normal places are considered when they are combined in the second step of the procedure.

As previously remarked, the required factorization of $\boldsymbol{A}_j$ is possible if its rank is at most p, the dimension of the normal place. On the other hand, the calculation of the normal place through (8) requires that the rank of $\boldsymbol{B}_j$ is exactly p, so that we should have $\mathrm{rank}(\boldsymbol{A}_j) = p$. In particular, at least p observations are needed in each group ($m_j \geq p$). In the following sections we consider how this requirement can be somewhat relaxed.

4 Using information arrays instead of normal places

The condition that $\boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j$ should be non-singular can be relaxed if the normal place and its covariance is replaced by the slightly more general concept of an *information array*. The information array is simply a compact representation of the normal equations in their properly normalized form – so that the inverse of the normal matrix, if it exists, equals the covariance of the least-squares estimate. For example, the information array representing the direct solution (3) consists of the $n(n+3)/2$ non-redundant elements of the $n \times (n+1)$ matrix

$$\left[\, \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} \quad \boldsymbol{A}^{\mathrm{T}} \boldsymbol{z} \,\right] \tag{16}$$

For each group, the processing of the observation equations (7) gives the partial information array

$$\left[\, \boldsymbol{F}_j \quad \boldsymbol{f}_j \,\right] = \left[\, \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{B}_j \quad \boldsymbol{B}_j^{\mathrm{T}} \boldsymbol{z}_j \,\right] \tag{17}$$

which of course exists also when $\boldsymbol{F}_j$ is singular. From (14) we now find that the information array for $\hat{\boldsymbol{x}}$ can be obtained by a simple processing of the partial information arrays:

$$\left[\, \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} \quad \boldsymbol{A}^{\mathrm{T}} \boldsymbol{z} \,\right] = \left[\, \sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{F}_j \boldsymbol{C}_j \quad \sum_j \boldsymbol{C}_j^{\mathrm{T}} \boldsymbol{f}_j \,\right] \tag{18}$$

This procedure not only avoids the difficulty with possibly singular solutions for the normal places, but is even simpler because the solutions need not be computed in the first place.

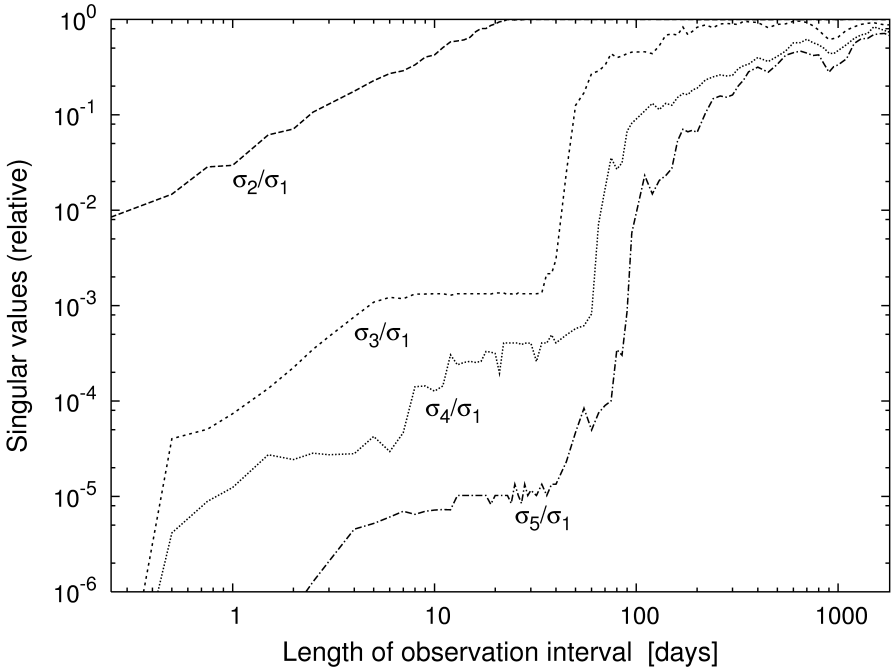

FIGURE 1: Evolution of the singular values of the partial observation matrices $\boldsymbol{A}_j$ for the five astrometric parameters, versus the length of the time interval confining the group j. Only one-dimensional (along-scan) observations are considered, and the plot shows the maximum relative singular value from many (≥ 1000) experiments using the nominal scanning law for random positions on the sky and for random starting times of the group.

5 Approximate decomposition

In reality it will seldom happen that the observation equations have a structure that will allow an exact decomposition according to (6).[1] Moreover, rounding errors might change the rank of the problem. Thus we must in practice consider the effective rank of the partial problems (4) and possibly allow some slight loss of accuracy due to the neglected components. Here, we should be guided by a good physical insight into the problem and numerical experiments.

As an example, let us consider the determination of the five astrometric parameters of a star from one-dimensional (along-scan) measurements made according to the nominal scanning law of Gaia. In this case the number of unknowns is $n = 5$, and the observations matrix obtained in a five-year mission will normally have full rank, $\text{rank}(\boldsymbol{A}) = n$. However, if a group of observations $\boldsymbol{A}_j$ is restricted to a short time interval, say at most T days, we will often find that $\text{rank}(\boldsymbol{A}) < n$. If we are going to use normal places of dimension $p < n$ to represent the observations in each group, it is necessary that $\text{rank}(\boldsymbol{A}_j) \leq p$ for every j. For given T, what is then the minimum p? Or, for given p, what is the maximum T?

[1] Disregarding the trivial cases $J = m$ (with $p = 1$) and $p = n$ (for any grouping).

The condition of each group is quantitatively given by the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ of $\boldsymbol{A}_j$, or more precisely by the relative singular values σ_i/σ_1 for $i = 2 \ldots n$ (assuming that each group contains at least one observation, so that $\sigma_1 > 0$). The interesting quantities are the maximum relative singular values, $\max_j(\sigma_i/\sigma_1)$, obtained for many different matrices $\boldsymbol{A}_j$, as function of the group interval length T. Figure 1 shows the result of such a calculation. The behaviour for the second largest singular value σ_2/σ_1 is roughly as expected: it shows that even for short time intervals of a day or less, you need to consider at least two-dimensional normal places. The behaviour for the next largest singular value σ_3/σ_1 is somewhat surprising. It remains below a level of $\simeq 10^{-3}$ up to $T \simeq 30$ days, after which it quickly rises.

A possible conclusion from this analysis is that, in the context of a Ring and Ring-to-Sphere type astrometric solution, one must use (at least) two-dimensional normal places in the Ring Solution, but that the time interval for each Ring Solution (using two-dimensional positions) could be as long as a month. Whether this is useful or practical from other considerations is of course a different matter.

Using two-dimensional normal places somehow seems to imply that the effects of proper motion and parallax are neglected, and therefore it is surprising that the actual sizes of the parallax or proper motion effects did not come into the discussion. After all, the proper motion over a 30-day interval will be very significant for most Gaia stars. But this is not a correct interpretation of what the normal places mean. They are always defined as estimates based on a group of residuals $(O–C)$ with respect to a reference solution (C), not of the observations (O) themselves. This is necessary in order to take into account the many other things that affect the observation, such as stellar aberration and calibration data. Eventually, the reference solution will have to include (most of) the proper motion and parallax effect as well, and their absolute sizes then become irrelevant for the solution. Thus iteration will be an essential feature of such a solution (also for other reasons), but its convergence properties might be very different from AGIS.

6 References

Bernstein, H., Bastian, U., Hirte, S., 2005, *First look - description of the ring solution*, GAIA-ARI-BST-001-5

ESA, 1997, *The Hipparcos and Tycho Catalogues*, ESA SP-1200

Hirte, S., Bernstein, H., Bastian, U., 2006, *Program description of the ring solution*, GAIA-C3-TN-ARI-SH-003

von Oppolzer, T.R., 1880, *Lehrbuch zur Bahnbestimmung der Kometen und Planeten. Zweiter Band*, Wilhelm Engelmann, Leipzig