Organization of PRS Experiments

L Lindegren 1983 March 18

The purpose of the experiments is to study the structure and properties of the PRS normals by small-scale simulations of 20–30 stars on the present mini-computer (HP 21MX) and perhaps 100 stars on the HP 9000. A major goal is to demonstrate the required rank deficiency due to the undefined coordinate system and to find a suitable solution method compatible with Cholesky. The small number of unknowns allows the full normals to be stored in primary memory. This is a considerable advantage with several algorithms e.g. for eigenvalue analysis, which are slow, requiring almost random access to matrix elements, and otherwise unsuitable for partitioning or sparse matrix methods.

The observation equations will use three kinds of unknowns ("parameters"):

1. Astrometric parameters: these are NA-vectors associated with each star ($2 \leq NA \leq 5$).

2. Set parameters: these are the abscissa zero points but possibly also other orientation or distortion parameters unique for each set. They are NB-vectors associated with each set ($1 \leq NB \leq 3$).

3. Global parameters: these form a single vector of length NG for the whole mission. They express a global, possibly time-dependent distortion of the coordinate system e.g. due to relativity or thermal effects on the instrument ($0 \leq NG \leq 10$).

The separation is motivated by the possibility to eliminate the second kind of unknowns at an early stage.

It is noted that this subdivision is exactly analogous to the one used for the set solutions, viz.

<table>
<thead>
<tr>
<th>PRS Solution</th>
<th>Set Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>astrometric parameters</td>
<td>abscissae</td>
</tr>
<tr>
<td>set parameters</td>
<td>frame attitudes ($\Delta \omega$)</td>
</tr>
<tr>
<td>global parameters</td>
<td>field distortion</td>
</tr>
</tbody>
</table>

With some foresight, it would be possible to use the same programme modules for forming and solving the normals in the two cases.
The development below is formulated in terms of the PRS solution, but exactly the same formulae apply for the set solutions, with only a different terminology. The notations are as follows (since "set" has a special meaning in our context, we use the term "class" for the general concept known from set theory):

\[ i = \text{star index, } i = 1, 2, \ldots, \text{NSTAR} \]
\[ j = \text{set index, } j = 1, 2, \ldots, \text{NSET} \]
\[ k = \text{observation index, } k = 1, 2, \ldots, \text{NOBS} \]
\[ i(k) = \text{star index for observation } k \]
\[ j(k) = \text{set index for observation } k \]
\[ k(i,j) = \text{the observation of star } i \text{ on set } j. \text{ This is assumed to be unique, i.e. at most one observation per star per set. Also we define} \]
\[ k(i,j) = 0 \text{ if star } i \text{ is not observed in set } j. \]
\[ K_j = \text{the class of observations constituting set } j: \{ k: j=j(k) \} \]
\[ K_i = \text{the class of observations concerned with star } i: \{ k: i=i(k) \} \]
\[ J_i = \text{the class of sets in which star } i \text{ is observed: } \{ j: k(i,j)\neq 0 \} \]
\[ I_j = \text{the class of stars observed in set } j: \{ i: k(i,j)\neq 0 \} \]
\[ \text{NSTAR} = \text{total number of stars} \]
\[ \text{NSET} = \text{total number of sets} \]
\[ \text{NOBS} = \text{total number of observations} \]
\[ \text{NA} = \text{number of astrometric parameters (per star), } 2 \leq \text{NA} \leq 5 \]
\[ \text{NB} = \text{number of set parameters (per set), } 1 \leq \text{NB} \leq 3 \]
\[ \text{NG} = \text{number of global parameters, } 0 \leq \text{NG} \leq 10 \]
\[ \text{NL} = \text{number of linear constraints (Lagrangian multipliers), } 0 \leq \text{NL} \leq 6 \]

The unknowns are:

\[ x_i, \ i = 1, 2, \ldots, \text{NSTAR}, \text{ an } (\text{NA}, 1)\text{-matrix for each star;} \]
\[ y_j, \ j = 1, 2, \ldots, \text{NSET}, \text{ an } (\text{NB}, 1)\text{-matrix for each set;} \]
\[ z, \text{ an } (\text{NG}, 1)\text{-matrix for the entire mission.} \]

After normalization (division by \( \sigma_k \), the m.e. of each observation), the observation equations are

\[ Ax + By + Cz = h \quad (1) \]

where \( A, B, \) and \( C \) are matrices of dimension \( (\text{NOBS}, \text{NSTAR} \times \text{NA}) \), \( (\text{NOBS}, \text{NSET} \times \text{NB}) \), and \( (\text{NOBS}, \text{NG}) \), respectively, and
\[
x = \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_{\text{NSTAR}}
\end{pmatrix} \quad y = \begin{pmatrix}
y_1 \\
y_2 \\
\vdots \\
y_{\text{NSET}}
\end{pmatrix}
\]

\( h \) is an \((\text{NOBS,1})\)-matrix containing the (normalized) \(G'-G's\).

The normals are

\[
A'Ax + A'By + A'Gz = A'h
\]
\[
B'Ax + B'By + B'Gz = B'h
\]  

\( G'Ax + G'By + G'Gz = G'h \) \hspace{1cm} (3)

Assuming that \(B'B\) is non-singular, we obtain by eliminating \(y\),

\[
A'DAx + A'DGz = A'Dh
\]
\[
G'DAx + G'DGz = G'Dh
\]  

\( \text{where } D \text{ is the symmetric idempotent } (\text{NOBS,NOBS})\)-matrix

\[
D = I - B(B'B)^{-1}B'
\]  

Elimination of \(y\) is thus effected through pre-multiplication by \(D\). Thanks to the block-diagonal form of \(B'B\) this operation can be performed for one set at a time, with only the observations of a single set in memory. The elements of \(B\) are the \((1,\text{NB})\)-matrices

\[
B_{kj} = b'_k \delta_{jj(k)}
\]  

with \(b_k\) being the NB-vector of coefficients for \(y_{j(k)}\). It is factored by

\[
\delta_{jj(k)} = 1 \text{ iff } j = j(k) \quad (\delta_{pq} \text{ is Kronecker's delta}).
\]

The elements of \(B'\) are

\[
(B')_{jk} = b'_k \delta_{jj(k)}
\]  

\( \text{whence}

\[
(B'B)_{jj} = \Sigma_k (B')_{jk} B_{kj} = \Sigma_k b'_k \delta_{jj(k)} b_k \delta_{j'-j(k)} = \delta_{jj} \Sigma_k b_k b'_k
\]  

\( \text{(8)} \)
The last equality follows from $\delta_{jj(k)} \delta_{j'}^j(k) = 1$ iff $j = j'$ and $k \in K_j$. We introduce the symmetric (NB,NB)-matrices

$$E_j = \sum_{k \in K_j} b_kb'_k$$

(9)

and assume that $E_j^{-1}$ exist. (Actually, if $E_j$ is singular for some $j$, the set does not contribute to the determination of the astrometric parameters and is simply skipped. This is analogous to skipping frames with less than two stars in the set solution. Thus we should understand summations over $j$ below to include only sets for which $E_j$ is non-singular.) Then

$$[(B'B)^{-1}]_{jj'} = \delta_{jj'} E_j^{-1}$$

(10)

and

$$D_{kk'} = I_{kk'} - [B(B'B)^{-1}B']_{kk'} =$$

$$= \delta_{kk'} - \sum_j \sum_{j} \sum_{j'} B_{kj} \delta_{jj'} E_j^{-1} (B')_{j'k'} =$$

$$= \delta_{kk'} - \delta_{j(k)j(k')} b'_{k} E_j^{-1} b_{k'}$$

(11)

Let $f$ be any vector of length NOBS. Pre-multiplying by $D$ gives a new vector $Df$ with elements

$$(Df)_k = f_k - b_k E_k^{-1} \sum_{k' \in K_j(k)} b_{k'} f_{k'}$$

(12)

The operation can be seen as subtracting a certain weighted mean of $f_k$ within each set. In the simplest case when $b_k$ is a constant scalar ($NB = 1$, $b_k = c_j$) we have for instance

$$(Df)_k = f_k - \frac{1}{n_j} \sum_{k' \in K_j(k)} f_{k'}$$

(13)

where $n_j =$ number of observations in set $j$. That $D$ is idempotent ($D^2 = D$) means that repeating the operation will not give a different vector.
The (partially reduced) normals (4) require the accumulation of $A'DA$, $A'DG$, $A'Dh$, $G'DG$, and $G'Dh$. When deriving algorithms for this, we need only take into account the special structure of $A$, since $G$ and $h$ are full matrices. The elements of $A$ are the $(1, NA)$-matrices

$$A_{ki} = a_k^i \delta_{ii(k)} \quad \text{(14)}$$

with $a_k^i$ being the NA-vector of coefficients for $x_i(k)$. Thus,

$$(A'DA)_{ii'} = \Sigma_k \Sigma_{k'} a_k^i \delta_{ii(k)} D_{kk'} a_{k'}^{i'} \delta_{ii'(k')} =$$

$$= \Sigma_k \Sigma_{k'} \delta_{ii(k)} \delta_{ii'(k')} \left[ \delta_{kk'} a_k^i a_{k'}^{i'} - \delta_{j(k)j(k')} a_k^j b_{k}^j E_{j(k)j(k')} a_{k'}^{i'} b_{k'}^{i'} \right] =$$

$$= \Sigma_{j \in J_i} a_k^i a_{k(j)}^{i(j)} - \Sigma_{j \in J_i \cap J_i} a_k^i a_{k(j)}^{i(j)} b_{k}^j E_{j(k)j(k')} b_{k'}^{i'} a_{k'}^{i'} a_{k'}^{i'}$$

$$\text{(15)}$$

The first sum is from the original (unreduced) normals, $(A'A)_{ii'}$, and is star-wise block-diagonal ($= 0$ for $i \neq i'$). The second sum is taken over sets including both $i$ and $i'$. In practice one takes of course one set at a time rather than one element $(ii')$, so that the algorithm would be as follows:

for each $j$

compute $E_j$; if singular go to next $j$, else compute $E_j^{-1}$

for each $k \in K_j$

$i = i(k)$

accumulate $a_k^i a_k^i$ to element $(i, i)$ of $A'DA$

for each $k' \in K_j$

$i' = i(k')$

if $i < i'$ go to next $k'$

compute the scalar $s = b_{k}^j E_{j(k)j(k')}^{-1} b_{k'}^{i'}$

accumulate $-a_k^i a_k^i$ to element $(i, i')$ of $A'DA$

next $k'$

next $k$

next $j$
Only the upper-triangular part of $A'\bar{D}A$ is computed ($i^- > i$) because of the symmetry. If the observations in a set are ordered after increasing $i(k)$, one could start the $k^-$-loop at $k$ instead of at the first observation of the set, but this is not recommended.

The formulae for $A'\bar{D}G$ and $A'\bar{D}h$ are similar to (15), except that the elimination of $y_j$ now establishes a link with all the stars in that set, not just with $i^-$ as in (15). Thus

$$
(A'\bar{D}G)_{11} = \sum_{j \in J_k} \left[ a_{k(i,j)}g_{k(i,j)} - a_{k(i,j)}b_{k(i,j)}^{-1} \sum_{i^- \in I_j} b_{k(i^-,j)}g_{k(i^-,j)} \right]
$$

(16)

if $g_k$ is the NG-vector of coefficients for $z$ in observation $k$:

$$
g'_{k1} = g_k
$$

(17)

The same equation (16) applies for $A'\bar{G}h$, with the elements of $h$, $h_k'$, replacing $g_k'$.

For the symmetric (NG,NG)-matrix $G'\bar{D}G$ we have

$$
(G'\bar{D}G)_{11} = \sum_k g_kg_k' - \sum_k g_kb_k'g_k^{-1} \sum_{k^- \in K_j(k)} b_{k^-}g_k'
$$

(18)

while for $G'\bar{D}h$ we replace $g_k'$ and $g_k''$ by $h_k'$ and $h_k''$.

The complete algorithm for accumulating the reduced normals is given on the next page.
Accumulation of reduced normals (A'DA, A'DG, A'Dh, G'DG, G'Dh)

for each $j$

compute $E_j$; if singular go to next $j$, else compute $E_j^{-1}$

for each $k \in K_j$

$i = i(k)$

accumulate $a_{k'k}, a_{k'k'}, a_{k'k'}, g_{k'k}, g_{k'k'}$ to respective elements

for each $k' \in K_j$

$i' = i(k')$

if $i' < i$ go to next $k'$

compute scalar $s = b_k^{-1} E_j b_{k'}$

accumulate $-a_{k'k'}, -a_{k'k'}, -a_{k'k'}, -g_{k'k}, -g_{k'k'}$

if $i' 
eq i$ accumulate also $-a_{k'k}, -a_{k'k}, -g_{k'k}$

next $k'$

next $k$

next $j$

Observe that the non-symmetric matrices A'DG, A'Dh and G'Dh must be accumulated for all combinations of $k$, $k'$, i.e. also when $i(k') < i(k)$. This is achieved in the line beginning ..if $i' 
eq i$., which takes advantage of that $s$ is the same after interchanging $k$ and $k'$.

Input file

The input file should be self-contained, including all data and parameters necessary for setting up, solving and analysing the normals. It is most convenient to define explicitly, at the beginning of the file, only the numbers NA, NB, NG, and NL, while NSTAR, NSET, and N0BS are implicitly defined by successive records. Sentinel records must be inserted to separate data groups. Since access is always sequential, file type 3 is preferred (variable-length records, any data type, extendable file length); the sentinel may then consist of a zero-length record. The first record (header) should contain only ASCII data. All subsequent records should contain unformatted data (memory binary representation) for compactness and to avoid truncation errors at formatting.
Tentatively, the file may be organized as follows (word = 16 bit):

<table>
<thead>
<tr>
<th>Record #</th>
<th>Length (words)</th>
<th>Content</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>ASCII header</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>NA, NB, NG, NL</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>4*NG</td>
<td>z, Δz</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1+4<em>NA+2</em>NL</td>
<td>i, x_i, Δx_i, u_i</td>
<td>1 record per star</td>
</tr>
<tr>
<td>3+NSTAR</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
</tr>
<tr>
<td>4+NSTAR</td>
<td>0</td>
<td>sentinel</td>
<td>end stars</td>
</tr>
<tr>
<td>5+NSTAR</td>
<td>1+4*NB</td>
<td>j, y_j, Δy_j</td>
<td>first set</td>
</tr>
<tr>
<td>6+NSTAR</td>
<td>2+2*(NA+NB+NG+2)</td>
<td>k, i, a_k, b_k, e_k, h_k, σ_k</td>
<td>n_1 obs.</td>
</tr>
<tr>
<td>5+NSTAR+n_1</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
</tr>
<tr>
<td>6+NSTAR+n_1</td>
<td>0</td>
<td>sentinel</td>
<td>end 1st set</td>
</tr>
<tr>
<td>7+NSTAR+n_1</td>
<td>1+4*NB</td>
<td>j, y_j, Δy_j</td>
<td>2nd set</td>
</tr>
<tr>
<td>8+NSTAR+n_1</td>
<td>2+2*(NA+NB+NG+2)</td>
<td>k, i, a_k, b_k, e_k, h_k, σ_k</td>
<td>n_2 obs.</td>
</tr>
<tr>
<td>7+NSTAR+n_1+n_2</td>
<td>&quot;</td>
<td>&quot;</td>
<td></td>
</tr>
<tr>
<td>8+NSTAR+n_1+n_2</td>
<td>0</td>
<td>sentinel</td>
<td>end 2nd set</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4+NSTAR+NOBS+2*NSET</td>
<td>0</td>
<td>sentinel</td>
<td>end last set</td>
</tr>
<tr>
<td>5+NSTAR+NOBS+2*NSET</td>
<td>0</td>
<td>sentinel (or EOF)</td>
<td>end file</td>
</tr>
</tbody>
</table>

Note: z is the vector of global parameters used for computing the O-C's (a priori values); Δz contains the corrections, z_true = z + Δz. Similarly for (x_i, Δx_i) and (y_j, Δy_j). u_i is an NL-vector containing the coefficients for the linear constraints. a_k, b_k, e_k and h_k are in general non-normalized σ_k ≠ 1.