JOURNAL OF RESEARCH of the National Bureau of Standards—D. Radio Propagation
Vol. 65D, No. 3, May–June 1961

Estimation of Variances of Position Lines From Fixes With Unknown Target Positions [1, 2]

E. M. L. Beale

(August 22, 1960; revised December 27, 1960)

Formulas are derived for the estimation of the variances of position lines from fixes with unknown target positions. Two approaches are considered, (1) that presented by Daniels [1951], and (2) an analysis of the squares of the errors in the position lines assuming the target is at the least squares estimate for its position.

1. Introduction

An important problem in position fixing is to estimate the accuracy of one's position lines. This can be done by assigning a variance to each line, see for example Daniels [1951] and Beale [1961]. These variances are usually estimated from fixes on targets whose true positions are known to the analyst. There are two major difficulties about estimating variances from fixes where the true position is unknown: (1) The estimates depend critically on the assumption that errors in different position lines are independent; and (2) The statistical problem is difficult, and any valid method seems to involve a substantial computing effort.

In spite of these difficulties, the problem deserves attention, because it is sometimes impossible to set up a satisfactory program of check fixes on known targets. Two approaches are considered in this paper. One follows Daniels [1951], the other is based on an analysis of the squares of the errors in the position lines assuming that the target is at the least squares estimate for its position.

Section 2 of this paper presents the basic assumptions common to both these approaches. The next 3 sections are concerned with Professor Daniels' approach to the problem. The approach is presented in general terms in section 3. Specific formulas for the important special case of 4 lines per fix are developed in section 4; and the application of these formulas is considered in section 5. The formulas for the alternative approach are developed in section 6, and their application is considered in section 7.

Finally, in section 8 some suggestions are made for an artificial sampling experiment to try out both these approaches.

2. Basic Assumptions

The basic assumptions made in this work are: (1) The earth is flat near the true position; (2) the position lines are straight lines; (3) an error of observation displaces the line parallel to itself; (4) the errors have zero means and are statistically independent; and (5) the errors are normally distributed, and we have rough estimates of their variances.

Given these assumptions, we can take Cartesian coordinates, and denote the j^{th} position line by the equation

$$x \sin \theta_j - y \cos \theta_j = p_j, \tag{2.1}$$

where θ_j is a known constant, and p_j is regarded as a random variable with mean

$$\xi \sin \theta_j - \eta \cos \theta_j,$$

[1] Contribution from Admiralty Research Laboratory, Teddington, Middlesex, England.

[2] Paper presented at the Conference on Transmission Problems Related to High-Frequency Direction Finding, at UCLA, June 21–24, 1960.

and variance σ_j^2, where (ξ,η) are the unknown coordinates of the true position. In our theoretical work, we assume that $\xi=\eta=0$, so that the mean value of p_j is zero; but we must remember that p_j is then not a directly observable quantity.

Some discussion of these basic assumptions may be helpful.

The first 3 assumptions are introduced to make the model linear, i.e., to make the value of each observation in the absence of errors a linear function of the unknown parameters ξ and η. It is clear that in some DF situations, notably when the DF stations have a narrow base line, the problem is decidedly nonlinear when expressed in terms of obvious parameters, such as the distances of the target east and north of some origin. But there is evidence to suggest that the "intrinsic nonlinearity" as defined by Beale [1960], is nearly always small for DF problems. This implies that any results derived from the standard linear theory will be valid, provided that they are expressed in terms independent of any particular system of parameters (and provided of course that any other assumptions made about the error distributions are valid).

Using the language of Beale [1960], we must define all the quantities we use in terms of sample space and the solution locus, and not solely in terms of parameter space (or possible target positions on the Earth, which is a representation of parameter space). Now a position line can be defined in sample space as the intersection with the solution locus of the hyperplane where one coordinate, i.e., one observation, is constant. All we require in this work is an interpretation of the quantities p_j, θ_j, and σ_j^2. This can be obtained by taking some specific point T on the Earth near the true target position, and imagining a parameter system that coincides locally with distances east and north of T. Now let β_j denote the actual j^{th} observation. (In the DF problem this is the bearing from the j^{th} DF station.) Let β_{jT} denote a hypothetical j^{th} observation giving a position line passing through T, and let $\rho_j(\beta_j)$ denote the (signed) distance on the Earth of the j^{th} position line from T as a function of β_j.

Let λ_j denote $d\,\rho_j/d\beta_j$, evaluated where $\beta_j=\beta_{jT}$.

Then we can write $p_j=\lambda_j(\beta_j-\beta_{jT})$, $\sigma_j^2=\lambda_j^2\text{var}\ (\beta_i)$, and θ_j is the angle that the hypothetical j^{th} position line passing through T makes with some arbitrary x-axis passing through T.

Note that it is not necessarily legitimate to interpret θ_j as the angle that the observed position line on the Earth makes with some x-axis; or to interpret λ_j as $d\rho_j/d\beta_j$ evaluated for the observed β_j. It is important that the same point T should be used for all position lines in the fix. (This approach can be used to derive the standard (Gauss) iterative procedure for finding the least squares estimate for the target position when the variances of the β_j are assumed known. The point T must then be taken as the trial value of the least squares point at each stage.)

Our 4th basic assumption, that the error distributions have zero means and are independent, is very important. In certain circumstances it will not be satisfied in practice, in which case the least squares point will usually be a less accurate estimate of the target's true position than standard theory suggests, even if the variances are correctly estimated; and furthermore the variances will be underestimated by the methods described in this paper. If we had numerical values for all correlations involved, we could allow for them, but in general there seems to be no practical alternative to assuming independence, hoping for the best, but realizing the possibility of being misled by correlated data.

Our 5th assumption, that the errors are normally distributed and we have rough estimates of their variances, is not so critical. It is only required to give appropriate weights in our least squares estimation of the variances, and to estimate the accuracy of our final variance estimates.

3. Daniels' Approach in General

The problem considered in this paper was considered by Daniels [1951]. In section 9, Daniels writes:

"When the true position corresponding to each cocked hat is unknown, . . . it is found that the method of maximum likelihood is useless for estimating the variances of the lines. It gives in general inconsistent answers, and sometimes no answer at all . . . . The method [breaks down because] the number of incidental parameters to be estimated increases with the number of cocked hats used.

"Nevertheless a simple method of estimating the variances does exist requiring no assumptions about the error distributions, provided parallel displacement is assumed and the errors are independent. Suppose there are N three-line cocked hats, not necessarily for the same true position, the lines having variances σ_1^2, σ_2^2, σ_3^2 to be determined. The expression

$$u=p_1 \sin(\theta_2-\theta_3)+p_2 \sin(\theta_3-\theta_1)+p_3 \sin(\theta_1-\theta_2) \tag{3.1}$$

has the same value whatever the point from which the perpendiculars p_1, p_2, and p_3 are dropped on to the lines, and u is easily measured by taking the point at a vertex. . . . If the point is chosen at the unknown true position it follows that

$$E(u^2)=\sigma_1^2 \sin^2(\theta_2-\theta_3)+\sigma_2^2 \sin^2(\theta_3-\theta_1)+\sigma_3^2 \sin^2(\theta_1-\theta_2). \tag{3.2}$$

If $N \geq 3$, the N equations with the observed values of u^2 replacing $E(u^2)$ may be solved by least squares. Since for normal errors var $(u^2)=2\{E(u^2)\}^2$, arbitrary weights have to be used in the first solution and correct weights approximated to in subsequent iterations. The estimates are, however, unbiased for any error distribution whatever the weights used."

But it should be noted that with fixed DF stations the normal equations derived from (3.2) to estimate the variances are liable to be very ill conditioned, and may even be singular. If the simplifying assumption about parallel displacements were strictly correct, then the values of θ_1, θ_2, and θ_3 would be the same for all tasks, so that the right hand sides of all equations of the form (3.2) would be strictly proportional, and one could never hope to estimate more than the given linear function of σ_1^2, σ_2^2, and σ_3^2. One might hope in practice to be saved by the error in the approximation; since θ_j may be effectively constant for all reasonably possible observations on a given target, but not constant over all targets. But this will only be so if the targets are well distributed about the DF stations.

It is fairly easy to see that this difficulty is fundamental to the problem, and is not simply a defect in the present approach. For if the θ_j were strictly constant, one could for example assume that $\sigma_1=\sigma_2=0$, and the signed distance of the point of intersection of the first two position lines from the third will have a certain probability distribution (normal if the bearing errors are normal) with mean zero and variance given by (3.2). This variance is constant as long as the true values of σ_1^2, σ_2^2, and σ_3^2 are constant. So data of this type can never disprove the hypothesis that $\sigma_1=\sigma_2=0$.

In the analogous one-dimensional problem, if instruments 1 and 2 give independent observations x_1 and x_2 of an unknown scalar quantity ξ, and x_i has mean ξ and variance σ_i^2, then $E(x_1-x_2)^2=\sigma_1^2+\sigma_2^2$. With an arbitrarily large number of pairs of observations one can therefore get an arbitrarily accurate estimate of $\sigma_1^2+\sigma_2^2$, but cannot estimate σ_1^2 or σ_2^2 separately. But if a third independent instrument is available, then one can estimate $\sigma_1^2+\sigma_2^2$ by $(x_1-x_2)^2$, $\sigma_1^2+\sigma_3^2$ by $(x_1-x_3)^2$, and $\sigma_2^2+\sigma_3^2$ by $(x_2-x_3)^2$. Hence σ_1^2 is estimated by the mean value of $\frac{1}{2}(x_1-x_2)^2+\frac{1}{2}(x_1-x_3)^2-\frac{1}{2}(x_2-x_3)^2$, i.e., $x_1^2-x_1x_2-x_1x_3+x_2x_3$, as pointed out by Pearson [1902].

If one adds a fourth position line to the 2-dimensional problem, then one can obtain four expressions analogous to (3.2) by taking each set of three lines in turn. It turns out that these four expressions do not suffice to estimate the four variances σ_1^2, σ_2^2, σ_3^2, and σ_4^2; because the determinant of the left hand side of the four equations vanishes identically. But Professor Daniels has pointed out in a private communication that one can get more expressions by considering the mean values of the products of u-statistics for different triangles. Thus if

$$u_{123}=p_1 \sin(\theta_2-\theta_3)+p_2 \sin(\theta_3-\theta_1)+p_3 \sin(\theta_1-\theta_2),$$

and

$$u_{124}=p_1 \sin(\theta_2-\theta_4)+p_2 \sin(\theta_4-\theta_1)+p_4 \sin(\theta_1-\theta_2),$$

then

$$E(u_{123}u_{124})=\sigma_1{}^2 \sin(\theta_2-\theta_3) \sin(\theta_2-\theta_4)+\sigma_2^2 \sin(\theta_3-\theta_1) \sin(\theta_4-\theta_1).$$

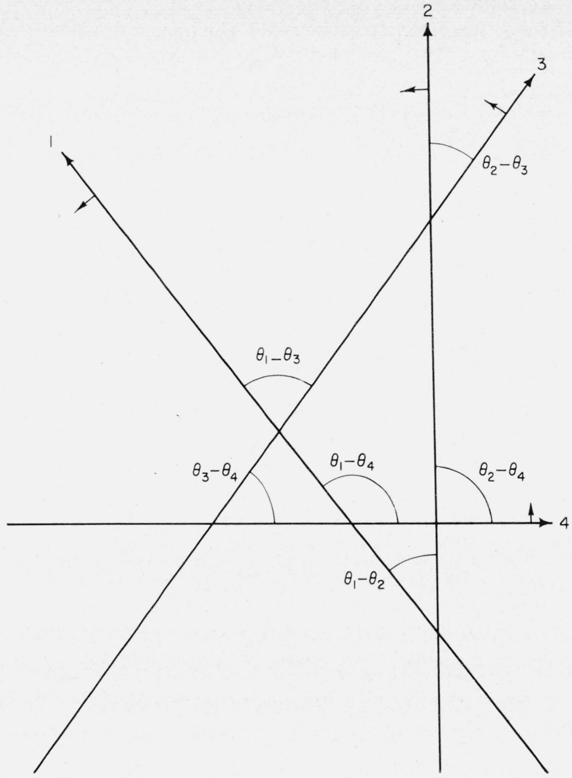

FIGURE 1. *A possible set of 4 position lines.*

Attention must be paid to signs. It is convenient to regard each position line as a directed line, pointing in the direction θ from some apex at an arbitrarily large distance from the scene of action. One can then define the corresponding displacement p as positive for points whose bearing from the apex is greater than θ (by an arbitrarily small amount) and negative on the other side of the position line. This will ensure that the signs of the terms on the right hand sides of the expressions analogous to (3.1) come out as expected. But it will still be necessary to study the configuration of lines to see which u-statistics are positive and which are negative. In the situation illustrated in figure 1, u_{134} and u_{234} are positive, and u_{123} and u_{124} are negative.

Given fixes with $n>3$ position lines per fix, one can obtain $U=n(n\text{-}1)(n\text{-}2)/6$ u-statistics from each fix, and $U(U+1)/2$ derived statistics by considering the squares and products of the u-statistics.

If $n=4$, this gives 4 u-statistics and 10 derived statistics.

If $n=5$, we have 10 u-statistics and 55 derived statistics.

This approach is therefore not very practical in its present form for $n\geq 5$. But the case $n=4$ is important, as it is the smallest n for which unique variance estimates can be obtained from a set of fixes, each with essentially the same values of θ_j, though not all on the same target. We therefore explore this case in more detail in the next section.

4. Daniels' Approach With $n=4$

With 4 position lines one can form 4 triangles, and obtain 4 basic quantities u_{123}, u_{124}, u_{134}, and u_{234}, which have the following expressions

$$\left.\begin{aligned} u_{123}&=p_1s_{23}+p_2s_{31}+p_3s_{12} \\ u_{124}&=p_1s_{24}+p_2s_{41}\qquad\quad +p_4s_{12} \\ u_{134}&=p_1s_{34}\qquad\quad +p_3s_{41}+p_4s_{13} \\ u_{234}&=\qquad\quad p_2s_{34}+p_3s_{42}+p_4s_{23}, \end{aligned}\right\} \tag{4.1}$$

where s_{ij} denotes sin $(\theta_i-\theta_j)$.

From these one can derive a vector $\underset{\sim}{\mathrm{y}}$ of 10 derived observations, such that

$$\underset{\sim}{\mathrm{y}}=\underset{\sim}{\mathrm{A}}\underset{\sim}{\mathrm{q}}, \tag{4.2}$$

where

$$\left.\begin{aligned} &y_1=u_{123}^2,\ y_2=u_{123}u_{124},\ y_3=u_{123}u_{134},\ y_4=u_{123}u_{234}, \\ &y_5=u_{124}^2,\ y_6=u_{124}u_{134},\ y_7=u_{124}u_{234}, \\ &y_8=u_{134}^2,\ y_9=u_{134}u_{234}, \\ &y_{10}=u_{234}^2; \end{aligned}\right\} \tag{4.3}$$

$$\left.\begin{aligned} &q_1=p_1^2,\ q_2=p_1p_2,\ q_3=p_1p_3,\ q_4=p_1p_4, \\ &q_5=p_2^2,\ q_6=p_2p_3,\ q_7=p_2p_4, \\ &q_8=p_3^2,\ q_9=p_3p_4, \\ &q_{10}=p_4^2; \end{aligned}\right\} \tag{4.4}$$

and A is

$$\begin{bmatrix} s_{23}^2 & -2s_{13}s_{23} & 2s_{12}s_{23} & 0 & s_{13}^2 & -2s_{12}s_{13} & 0 & s_{12}^2 & 0 & 0 \\ s_{23}s_{24} & (-t_2-t_3) & s_{12}s_{24} & s_{12}s_{23} & s_{13}s_{14} & -s_{12}s_{14} & -s_{12}s_{13} & 0 & s^2{}_{12} & 0 \\ s_{23}s_{34} & -s_{13}s_{34} & (t_1-t_3) & s_{13}s_{23} & 0 & s_{13}s_{14} & -s_{13}^2 & -s_{12}s_{14} & s_{12}s_{13} & 0 \\ 0 & s_{23}s_{34} & -s_{23}s_{24} & s_{23}^2 & -s_{13}s_{34} & (t_1+t_2) & -s_{13}s_{23} & -s_{12}s_{24} & s_{12}s_{23} & 0 \\ s_{24}^2 & -2s_{14}s_{24} & 0 & 2s_{12}s_{24} & s_{14}^2 & 0 & -2s_{12}s_{14} & 0 & 0 & s_{12}^2 \\ s_{24}s_{34} & -s_{14}s_{34} & -s_{14}s_{24} & (t_1+t_2) & 0 & s_{14}^2 & -s_{13}s_{14} & 0 & -s_{12}s_{14} & s_{12}s_{13} \\ 0 & s_{24}s_{34} & -s_{24}^2 & s_{23}s_{24} & -s_{14}s_{34} & s_{14}s_{24} & (t_1-t_3) & 0 & -s_{12}s_{24} & s_{12}s_{23} \\ s_{34}^2 & 0 & -2s_{14}s_{34} & 2s_{13}s_{34} & 0 & 0 & 0 & s_{14}^2 & -2s_{13}s_{14} & s_{13}^2 \\ 0 & s_{34}^2 & -s_{24}s_{34} & s_{23}s_{34} & 0 & -s_{14}s_{34} & s_{13}s_{34} & s_{14}s_{24} & (-t_2-t_3) & s_{13}s_{23} \\ 0 & 0 & 0 & 0 & s_{34}^2 & -2s_{24}s_{34} & 2s_{23}s_{34} & s_{24}^2 & -2s_{23}s_{24} & s_{23}^2 \end{bmatrix} \tag{4.5}$$

where $t_1=s_{12}s_{34}$, $t_2=s_{13}s_{24}$ and $t_3=s_{14}s_{23}$.

Then

$$E(\underset{\sim}{\mathrm{y}})=\underset{\sim}{\mathrm{H}}(\sigma_1{}^2,\ \sigma_2^2,\ \sigma_3^2,\ \sigma_4^2)', \tag{4.6}$$

where $\underset{\sim}{H}$ is the matrix formed from the 1st, 5th, 8th, and 10th colunms of $\underset{\sim}{\mathrm{A}}$, since the p_j are independent and have zero means when referred to the true target position as origin.

But in order to determine rational, as opposed to arbitrary, least squares estimates for the unknown parameters σ_1^2, σ_2^2, σ_3^2, and σ_4^2, it is necessary to consider the covariance matrix of $\underset{\sim}{\mathrm{y}}$, We rely on the fact that the p_j are independent and have zero means, and find that

$$c_{ij}=\mathrm{cov}(y_iy_j)=Ey_iy_j-(Ey_i)(Ey_j)=\sum_k a_{ik}a_{jk}r_k, \tag{4.7}$$

where

$$\left.\begin{aligned} &r_1=2\sigma_1^4,\ r_2=\sigma_1^2\sigma_2^2,\ r_3=\sigma_1^2\sigma_3^2,\ r_4=\sigma_1^2\sigma_4^2, \\ &r_5=2\sigma_2^4,\ r_6=\sigma_2^2\sigma_3^2,\ r_2=\sigma_2^2\sigma_4^2, \\ &r_8=2\sigma_3^4,\ r_9=\sigma_3^2\sigma_4^2, \\ &r_{10}=2\sigma_4^4. \end{aligned}\right\} \tag{4.8}$$

(The coefficients 2 in r_1, r_5, r_8, and r_{10} should be replaced by $2+\kappa_4$ if the error distributions have a fourth cumulant κ_4 different from zero.)

The covariance matrix $\underset{\sim}{C}=(c_{ij})$ of the derived observations is therefore obtained from (4.7), or equivalently by defining a new matrix

$$\underset{\sim}{F}=(f_{ij}),$$

$$\text{where } f_{ij}=a_{ij}r_j^{1/2},$$

$$\text{so that } \underset{\sim}{C}=\underset{\sim}{F}\,\underset{\sim}{F}'.$$

Now the standard least squares results can be expressed in matrix notation as follows: If $\underset{\sim}{Y}=\underset{\sim}{X}\underset{\sim}{\beta}+\underset{\sim}{\epsilon}$, where $\underset{\sim}{Y}$ is a vector of observations, $\underset{\sim}{X}$ is a known matrix, $\underset{\sim}{\beta}$ is a vector of unknown parameters, and $\underset{\sim}{\epsilon}$ is a random vector whose components are uncorrelated and have means zero and a common variance σ^2, then the least squares estimates $\underset{\sim}{B}$ of $\underset{\sim}{\beta}$ are given by choosing $\underset{\sim}{B}$ to minimize

$$Q=(\underset{\sim}{X}\underset{\sim}{B}-\underset{\sim}{Y})'(\underset{\sim}{X}\underset{\sim}{B}-\underset{\sim}{Y}), \tag{4.9}$$

Assuming that the errors are normally distributed, the likelihood is proportional to $\exp\,(-Q/2\sigma^2)$, and in any case

$$\underset{\sim}{B}=(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{Y}, \tag{4.10}$$

and the covariance matrix for the estimates $\underset{\sim}{B}$ is given by $(\underset{\sim}{X}'\,\underset{\sim}{X})^{-1}\sigma^2$. (4.11)

If now $\underset{\sim}{y}=\underset{\sim}{G}\underset{\sim}{Y}$, we have $\underset{\sim}{y}=\underset{\sim}{G}\underset{\sim}{X}\underset{\sim}{\beta}+\underset{\sim}{G}\underset{\sim}{\epsilon}$, and therefore the covariance of y_i and y_j is given by

$$\sum_k g_{ik}g_{jk}\sigma^2.$$

Hence we can identify $\underset{\sim}{G}$ with the matrix $\underset{\sim}{F}$, putting $\sigma^2=1$, and we know that, if $\beta=(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2)'$,

$$\underset{\sim}{y}=\underset{\sim}{H}\underset{\sim}{\beta}+\underset{\sim}{F}\epsilon.$$

So $\underset{\sim}{X}=\underset{\sim}{F}^{-1}\underset{\sim}{H}$, $\underset{\sim}{Y}=\underset{\sim}{F}^{-1}\underset{\sim}{y}$,

and

$$\begin{aligned} Q&=\{\underset{\sim}{F}^{-1}(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})\}'\{\underset{\sim}{F}^{-1}(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})\}\\ &=(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})'\underset{\sim}{F}^{-1\prime}\underset{\sim}{F}^{-1}(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})\\ &=(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})'(\underset{\sim}{F}\underset{\sim}{F}')^{-1}(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})\\ &=(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y})'\underset{\sim}{C}^{-1}(\underset{\sim}{H}\underset{\sim}{\beta}-\underset{\sim}{y}). \end{aligned} \tag{4.12}$$

Further

$$\begin{aligned} \underset{\sim}{\beta}&=(\underset{\sim}{H}'\underset{\sim}{F}'^{-1}\underset{\sim}{F}^{-1}\underset{\sim}{H})^{-1}\underset{\sim}{H}'\underset{\sim}{F}'^{-1}\underset{\sim}{F}^{-1}\underset{\sim}{y}\\ &=(\underset{\sim}{H}'\underset{\sim}{C}^{-1}\underset{\sim}{H})^{-1}\underset{\sim}{H}'\underset{\sim}{C}^{-1}\underset{\sim}{y}, \end{aligned} \tag{4.13}$$

and the covariance matrix of $\underset{\sim}{\beta}$ is given by

$$(\underset{\sim}{H}'\underset{\sim}{C}^{-1}\underset{\sim}{H})^{-1}. \tag{4.14}$$

These results are due to Aitken [1935].

5. Application of the Formulas of Section 4

To apply these formulas to a situation in which one has N fixes, i.e., sets of four position lines, each with a constant variance and a constant value of θ_j, one proceeds as follows.

One first sorts out one's sign convention, as indicated at the end of section 3. This will produce signed values of u_{123}, u_{124}, u_{134}, and u_{234} for each of the N fixes. From these one computes derived observations $y_1, \ldots, y_{10}$ for each fix from (4.3). Averaging over all fixes, one obtains a vector

$$\underset{\sim}{\bar{\mathrm{y}}}=(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{10})'.$$

We next compute the matrix $\underset{\sim}{\mathrm{A}}$ from (4.5).

By using the guessed values for the unknown variances σ_1^2, σ_2^2, σ_3^2, and σ_4^2, one can derive values of $r_1^2, r_2^2, \ldots, r_{10}^2$ from (4.8), and hence compute the elements c_{ij} of the matrix $\underset{\sim}{\mathrm{C}}$ from (4.7).

This matrix must be inverted to produce $\underset{\sim}{\mathrm{C}}^{-1}$.

We must also consider the matrix $\underset{\sim}{\mathrm{H}}$ whose columns are the 1st, 5th, 8th, and 10th columns of $\underset{\sim}{\mathrm{A}}$, and our estimates of σ_1^2, σ_2^2, σ_3^2, and σ_4^2 are the 4 elements of the matrix

$$(\underset{\sim}{\mathrm{H}}'\underset{\sim}{\mathrm{C}}^{-1}\underset{\sim}{\mathrm{H}})^{-1}\underset{\sim}{\mathrm{H}}'\underset{\sim}{\mathrm{C}}^{-1}\underset{\sim}{\bar{\mathrm{y}}}. \tag{5.1}$$

If our guessed values of the σ_i^2, used to define $\underset{\sim}{\mathrm{r}}$, were correct, the covariance matrix for these estimates would be

$$\frac{1}{N}(\underset{\sim}{\mathrm{H}}'\underset{\sim}{\mathrm{C}}^{-1}\underset{\sim}{\mathrm{H}})^{-1}. \tag{5.2}$$

One could try iterating this procedure, using the estimated variances to produce a new vector $\underset{\sim}{\mathrm{r}}$, and hence a new $\underset{\sim}{\mathrm{C}}$, and hence revised estimated variances, and so on.

I am doubtful of the wisdom of this, though it might be interesting to try it on an artificial sampling experiment. If one wants a reasonable estimate of the accuracy of the estimates one must compute a revised $\underset{\sim}{\mathrm{C}}$ to use in (5.2). But I can see little merit and some danger in revising the weights to produce revised estimates: Note that the iterated estimates are not unbiased.

This procedure may on occasion produce negative estimated variances. Steps must be taken at least to ensure that negative (or even very small positive) variances are not used in the definition of $\underset{\sim}{\mathrm{r}}$. Theoretically the least squares problem should be set up as a quadratic programing problem when such negative estimates are produced by the standard method. But this is too large an issue to face at this stage.

6. A Direct Approach

In this section we derive some equations concerning the apparent error in the position lines, assuming that the target is at the least squares estimate for its position.

Suppose that the j^{th} position line is

$$x \sin \theta_j - y \cos \theta_j = p_j,$$

with the origin at the (unknown) true target position. Suppose further that we have some rough estimate s_j^2 for σ_j^2.

We write

$$\sigma_j^2 = s_j^2({}_j^2 1+\delta_j), \text{ and } 1/s_j^2 = w_j.$$

Now the sum of squares function, whose minimum value defines the least squares estimate, is

$$Q=\sum_j w_j(x \sin \theta_j - y \cos \theta_j - p_j)^2 = a\, x^2+2hxy+by^2+2gx+2fy+c,$$

where

$$\left.\begin{aligned} &a=\sum_j w_j \sin^2 \theta_j, \quad b=\sum_j w_j \cos^2 \theta_j, \quad c=\sum_j w_j p_j^2, \\ &f=\sum_j w_j p_j \cos \theta_j, \quad g=-\sum w_j p_j \sin \theta_j, \quad h=-\sum_j w_j \sin \theta_j \cos \theta_j. \end{aligned}\right\} \tag{6.1}$$

Note that a, b, and h are known constants, while f, g, and c depend on the random variables p_j.

It is convenient to rotate axes, so that in the new coordinates (X, Y),

$$Q=AX^2+BY^2+2GX+2FY+C.$$

This involves finding an angle α such that

$$\phi_j=\theta_j+\alpha, \tag{6.2}$$

and the jth position line is

$$X \sin \phi_j - Y \cos \phi_j = p_j,$$

i.e.,

$$X \sin (\theta_j+\alpha) - Y \cos (\theta_j+\alpha) = p_j,$$

so that

$$x = X \cos \alpha + Y \sin \alpha,$$

$$y = -X \sin \alpha + Y \cos \alpha.$$

And since $ax^2+2hxy+by^2+2gx+2fy+c=AX^2+2HXY+BY^2+2GX+2FY+C$, where $H=0$, we have

$$\left.\begin{aligned}
A&=a \cos^2 \alpha - 2h \cos \alpha \sin \alpha + b \sin^2 \alpha,\\
0&=a \cos \alpha \sin \alpha + h(\cos^2 \alpha - \sin^2 \alpha) - b \cos \alpha \sin \alpha,\\
B&=a \sin^2 \alpha + 2h \cos \alpha \sin \alpha + b \cos^2 \alpha,\\
G&=g \cos \alpha - f \sin \alpha,\\
F&=g \sin \alpha + f \cos \alpha,\\
C&=c.
\end{aligned}\right\} \tag{6.3}$$

This implies that

$$\tan 2\alpha = 2h/(b-a). \tag{6.4}$$

This defines a set of values of α differing by multiples of 90°. It does not matter which is taken.

The values of A, B, C, F, and G can be obtained either from (6.1) and (6.3), or alternatively from the formulas

$$\left.\begin{aligned}
&A=\sum_j w_j \sin^2 \phi_j, \quad B=\sum_j w_j \cos^2 \phi_j, \quad C=\sum_j w_j p_j^2,\\
&F=\sum_j w_j p_j \cos \phi_j, \quad G=-\sum_j w_j p_j \sin \phi_j, \quad H=-\sum_j w_j \sin \phi_j \cos \phi_j = 0,
\end{aligned}\right\} \tag{6.5}$$

where ϕ_j is obtained by substituting for α from (6.4) in (6.2).

The transformed coordinates of the least squares estimate are then

$$X_0=-G/A \text{ and } Y_0=-F/B. \tag{6.6}$$

Now let d_k denote the signed displacement of the k^{th} position line from the least squares estimate. This quantity can be observed in practice. Then

$$d_k = p_k + \frac{G}{A} \sin \phi_k - \frac{F}{B} \cos \phi_k = \sum_{i=1}^{n} p_i (w_i/w_k)^{1/2} \lambda_{ik}, \tag{6.7}$$

where

$$\lambda_{kk} = 1 - w_k \left(\frac{\sin^2 \phi_k}{A} + \frac{\cos^2 \phi_k}{B} \right), \tag{6.8}$$

$$\lambda_{ik} = \lambda_{ki} = -(w_i w_k)^{1/2} \left(\frac{\sin \phi_i \sin \phi_k}{A} + \frac{\cos \phi_i \cos \phi_k}{B} \right), \ (i \neq k). \tag{6.9}$$

Note that

$$\sum_{k=1}^{n} \lambda_{kk} = n - \sum w_k \sin^2 \phi_k / A - \sum w_k \cos^2 \phi_k / B = n-2. \tag{6.10}$$

Further

$$\sum_{i=1}^{n}\lambda_{ik}^2=1-2w_k\left(\frac{\sin^2\phi_k}{A}+\frac{\cos^2\phi_k}{B}\right)+w_k\sum_{i=1}^{n}\left\{\frac{w_i\sin^2\phi_i\sin^2\phi_k}{A^2}+2\frac{w_i\sin\phi_i\cos\phi_i\sin\phi_k\cos\phi_k}{AB}+\frac{w_i\cos^2\phi_i\cos^2\phi_k}{B^2}\right\}=1-w_k\left(\frac{\sin^2\phi_k}{A}+\frac{\cos^2\phi_k}{B}\right).$$

So

$$\sum_{i=1}^{n}\lambda_{ik}^2=\lambda_{kk}. \tag{6.11}$$

Further, if $k\neq l$,

$$\sum_{i=1}^{n}\lambda_{ik}\lambda_{il}=2\lambda_{kl}+(w_kw_l)^{1/2}\sum_{i=1}^{n}w_i\left(\frac{\sin\phi_i\sin\phi_k}{A}+\frac{\cos\phi_i\cos\phi_k}{B}\right)\left(\frac{\sin\phi_i\sin\phi_l}{A}+\frac{\cos\phi_i\cos\phi_l}{B}\right)$$

$$=2\lambda_{kl}+(w_kw_l)^{1/2}\left(\frac{\sin\phi_k\sin\phi_l}{A}+\frac{\cos\phi_k\cos\phi_l}{B}\right)\text{ from (6.5)}=\lambda_{kl}.$$

So (6.11) can be generalized to read

$$\sum_{i=1}^{n}\lambda_{ik}\lambda_{il}=\lambda_{kl}. \tag{6.12}$$

From (6.7) we deduce that

$$w_kd_k^2=\sum_{i=1}^{n}\sum_{j=1}^{n}(w_iw_j)^{1/2}\lambda_{ik}\lambda_{jk}p_ip_j. \tag{6.13}$$

Now

$$Ep_i^2=\sigma_i^2=(1+\delta_i)/w_i,\ Ep_ip_j=0 \qquad (i\neq j),$$

So

$$Ew_kd_k^2=\sum_{i=1}^{n}\lambda_{ik}^2(1+\delta_i)=\lambda_{kk}+\sum_{i=1}^{n}\delta_i\lambda_{ik}^2, \qquad \text{from (6.11)}. \tag{6.14}$$

It follows that

$$E\sum_{k=1}^{n}w_kd_k^2=\sum_{k=1}^{n}\lambda_{kk}+\sum_{i=1}^{n}\delta_i\sum_{k=1}^{n}\lambda_{ik}^2=n-2+\sum_{i=1}^{n}\delta_i\lambda_{ii}, \qquad \text{from (6.10) and (6.11)}. \tag{6.15}$$

If $\delta_i=\delta$ for all i, we have the well-known formula

$$E\sum w_kd_k^2=(n-2)(1+\delta), \tag{6.16}$$

(which can easily be deduced from first principles).

It is also of interest to consider the covariance of $w_kd_k^2$ and $w_ld_l^2$.

We have

$$\operatorname{cov}(w_kd_k^2,w_ld_l^2)=\sum_{i=1}^{n}w_i^2\lambda_{ik}^2\lambda_{il}^2 2(1+\delta_i)^2/w_i^2$$

$$+2\sum_{i=1}^{n}\sum_{j\neq i}w_iw_j\lambda_{ik}\lambda_{jk}\lambda_{il}\lambda_{jl}(1+\delta_i)(1+\delta_j)/w_iw_j \qquad \text{by analogy with (4.7)}$$

$$=2\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{ik}\lambda_{jk}\lambda_{il}\lambda_{jl}(1+\delta_i)(1+\delta_j). \tag{6.17}$$

It is of particular interest to consider the covariance matrix when all $\delta_i=0$, since this corresponds to the originally estimated covariance matrix.

Using (6.12) we find this reduces to

$$\operatorname{cov}(w_kd_k^2,w_ld_l^2)=2\lambda_{kl}^2. \tag{6.18}$$

If $n=3$, it can be shown that $\lambda_{11}/\lambda_{21}=\lambda_{12}/\lambda_{22}=\lambda_{13}/\lambda_{23}$, etc. For example, to prove that $\lambda_{11}\lambda_{22}-\lambda_{12}^2=0$, one substitutes for the λ_{ij} from (6.8) and (6.9) and multiplies out. The terms in $1/A^2$ and $1/B^2$ cancel, so the expression can be written as a quadratic function of w_1,w_2 and w_3 (using the definitions of A and B) divided by AB, and H proves to be a factor of the numerator.

This is an algebraic proof of something we already know on statistical grounds: that with only 3 bearings the random variables $w_1d_1^2$, $w_2d_2^2$ and $w_3d_3^2$ are totally correlated, and it is impossible to estimate more than one linear function of the unknown variances.

7. Application of the Formulas of Section 6

To apply the formulas of section 6 to a situation in which one has N fixes, i.e., sets of n position lines, each with a constant variance and a constant value of θ_j, one proceeds as follows.

One starts with some assumed variances s_j^2 for the position lines, and one uses those to compute a least squares point estimate for the target for each fix. One then computes $w_jd_j^2$, i.e., d_j^2/s_j^2, being the weighted square of the displacement of the j^{th} position line from each fix.

One then averages over all N fixes, obtaining n quantities $\overline{w_1d_1^2}, \ldots, \overline{w_nd_n^2}$.

One must also compute λ_{ii} for $i=1, \ldots, n$, and λ_{ij}^2 for $i, j=1, \ldots, n$. These quantities are the same for all fixes.

Note that λ_{ii} is defined by (6.8) and λ_{ij} by (6.9) if $i\neq j$.

It may be more convenient to use the formula

$$\lambda_{ij}^2=\frac{(w_i\sin^2\phi_i)(w_j\sin^2\phi_j)}{A^2}+\frac{2(w_i\sin\phi_i\cos\phi_i)(w_j\sin\phi_j\cos\phi_j)}{AB}+\frac{(w_i\cos^2\phi_i)(w_j\cos^2\phi_j)}{B^2},\quad (i\neq j). \tag{7.1}$$

Then, from (6.14), we solve the equations

$$\sum_{i=1}^{n}\hat{\delta}_i\lambda_{ik}^2=\overline{w_kd_k^2}-\lambda_{kk},\ (k=1, \ldots, n). \tag{7.2}$$

and the resulting values of $\hat{\delta}_i$ define the estimated variances, since

$$\sigma_i^2\approx s_i^2(1+\hat{\delta}_i).$$

To find the estimated covariance matrix for these estimates $\hat{\delta}_i$, we form the matrix $\underset{\sim}{H}$ such that

$$(h_{ij})=\lambda_{ij}^2,$$

and the matrix $\underset{\sim}{C}$ such that

$$(c_{ij})=2\lambda_{ij}^2\quad \text{(from (6.18))}. \tag{7.3}$$

The estimated covariance matrix is then given by (5.2), if the $\hat{\delta}_i$ prove to be small, and this reduces to $\frac{4}{N}\underset{\sim}{C}^{-1}$.

If one wants to combine these data with other data, then the contribution of these data to the sum of squares to be minimized is then NQ, where Q is given by (4.12) with $\underset{\sim}{\beta}=(\delta_1, \ldots, \delta_n)'$ and $\underset{\sim}{y}=(\overline{w_1d_1^2}, \ldots, \overline{w_nd_n^2})'$.

8. Proposed Computational Program

Some computational experience with both the schemes proposed here would be very valuable.

In the first instance, it would be of interest to inspect the covariance matrix (4.14) for estimates obtained by Daniels' approach, i.e., with $\underset{\sim}{H}$ defined following (4.6) and C defined by (4.7), for various values of θ_j and σ_j^2.

It would also be of interest to inspect the corresponding matrices for the direct approach, i.e., with $(h_{ij})=(\lambda_{ij}^2)$ and $(c_{ij})=(2\lambda_{ij}^2)$. In this case (4.14) reduces to 4 $\underset{\sim}{C}^{-1}$. The $(ij)^{\text{th}}$ element of this covariance matrix for the $\hat{\delta}$ must be course of multiplied by $s_i^2 s_j^2$ to represent the covariance matrix for the variances. It will then be directly comparable with the other.

One may hope to throw considerable further light on these procedures by an artificial sampling experiment. One must first choose values of θ_j, σ_j^2, and (for the direct approach) s_j^2. These will define values of a, b, and h from (6.1), and hence of A, B and α from (6.3) and (6.4). One can then compute the λ_{ik} from (6.9).

One then takes sets of values of p_i as pseudo-random normal deviates with means zero and variances σ_i^2, and use these to generate synthetic u-statistics from (4.1), and values of d_i from (6.7). One can then compare the estimated variances obtained by the two approaches (a) after one iteration, and (b) after using the results of each iteration as starting values for the next (i.e., to compute the r_i for Daniels' approach and the w_i for the direct approach) and continuing until the variances that go in come out.

One can also consider the effects on the two approaches of using poor starting approximations to the variances.

It seems likely on general grounds that Daniels' approach should give better results, since it uses more information. But if the direct approach is at all satisfactory with $n=4$, one may hope that it will be still better with larger values of n.

9. References

Aitken, A. C., On least squares and linear combination of observations, Proc. Roy. Soc. Edinburgh **55,** 42 (1935).

Beale, E. M. L., Confidence regions in nonlinear estimation, J. Roy. Stat. Soc. (B) **22,** 41 (1960).

Beale, E. M. L., The Brooke variance classification system, J. Research NBS **65D**, No. 3, (May–June 1961).

Daniels, H. E., The theory of position finding, J. Roy. Stat. Soc. (B) **13,** 186 (1951).

Pearson, K., On the mathematical theory of errors of judgment, with special reference to the personal equation, Phil. Trans. Roy. Soc. London, Series A, **198,** 235 (1902).

(Paper 65D3–129)