H. S. WILLIAMS*
University of the Witwatersrand
Johannesburg, South Africa

A Theorem in Least Squares

Omission or addition of observations after parametric adjustment.

INTRODUCTION

ON SOME occasions in adjusting indirect observations it becomes necessary to amend the estimates of the unknowns as a result of (*i*) the omission of one, or more, doubtful observations, or (*ii*) the inclusion of additional new observations.

If the number of unknowns is small and automatic digital computing facilities are readily available, the quickest solution to either of the problems mentioned above is to run the entire adjustment again. However, if the system of equations is large and if the observations to be omitted, or added, are few in number, it is possible to amend existing results in a relatively simple way without having to repeat the inversion of the large matrix of coefficients of the normal equations system, $A^t w A$. The most important practical applications are where the observations for the solution amendment are one, two or three in number. A possible correction procedure is given below.

ABSTRACT: *Attention is drawn to a property of parameter adjustment (adjustment of indirect observations) which enables an estimate of the unknowns, based on* n ± k *observations, to be obtained once the adjustment for* n *observations has been calculated. If* k *is small the correction procedure is simple. An example is provided to illustrate the method.*

DERIVATION OF THE CORRECTION TERMS

Consider an $m \times 1$ vector of unknown parameters ΔX related to an $n \times 1$ vector of residuals v by the equation

$$A \cdot \Delta X + f = v \tag{1}$$

where A is an $n \times m$ matrix of known coefficients and f is another $n \times 1$ vector of known or absolute quantities. If we write

$$f = F_0 - l \tag{2}$$

$$L = l + v \tag{3}$$

and identify the l as the observations of the quantities L, the v are then recognized as the corrections to the l. F_0 will be seen to include approximate values, (X), of the unknowns, such that

$$X = (X) + \Delta X. \tag{4}$$

If (*i*) $n = m$ the solution for the unknowns ΔX is unique; (*ii*) if $n < m$, no unique solution; and (*iii*) if $n > m$ an overdetermination of the unknowns occurs. If w^{-1} is the $n \times n$ variance/covariance matrix of the observational errors, this will be a diagonal matrix where the observations of the quantities involved in the adjustment are treated as being independent.

* The article was published in *The South Africa Survey Journal* in December 1971.

The *method of least squares* leads to unbiased estimates ΔX^* of the unknowns ΔX with minimum variance which satisfy the normal equations

$$A^t w A \cdot \Delta X^* + A^t w f = 0. \tag{5}$$

That is,

$$\Delta X^* = -(A^t w A)^{-1} A^t w f \tag{6}$$

If v_2, of order $k \times 1$, is the vector of corrections to the observations to be omitted, Equation 1 may be rewritten appropriately as

$$v = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \Delta X + \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{7}$$

The submatrix A_1 is of order $(n-k) \times m$, such that $(n-k) \geq m$. Applied to Equation 7, the method of least squares provides the unbiased estimates of the unknowns as

$$\Delta X^* = -(A_1^t w_1 A_1 + A_2^t w_2 A_2)^{-1}(A_1^t w_1^t w_1 f_1 + A_2^t w_2 f_2). \tag{8}$$

Comparing Equations 6 and 8 we see that

$$A^t w A = A_1^t w_1 A_1 + A_2^t w_2 A_2$$

or

$$Q = Q_1 + Q_2,$$

and

$$A^t w f = A_1^t w_1 f_1 + A_2^t w_2 f_2$$

or

$$F = F_1 + F_2.$$

The first set of reduction formulas are then obtained simply as

$$Q_1 = Q - A_2^t w_2 A_2 \tag{9}$$

and

$$F_1 = F - A_2^t w_2 f_2 \tag{10}$$

Now $Q - Q_1 = A_2^t w_2 A_2$, which on premultiplication by $A_2 Q^{-1}$ and postmultiplication by Q_1^{-1} becomes

$$A_2 Q^{-1}(Q - Q_1) Q_1^{-1} = A_2 Q^{-1} A_2^t w_2 A_2 Q_1^{-1}.$$

Therefore,

$$A_2 Q_1^{-1} - A_2 Q^{-1}(Q - Q_1) Q_1^{-1} = A_2 Q_1^{-1} - A_2 Q^{-1} A_2^t w_2 A_2 Q_1^{-1}$$

and

$$A_2 Q^{-1} = (I - A_2 Q^{-1} A_2^t w_2) A_2 Q_1^{-1}.$$

Now, premultiplying both sides of this equation by $w_2(I - A_2 Q^{-1} A_2^t w_2)^{-1}$ we find that

$$w_2 A_2 Q_1^{-1} = w_2(I - A_2 Q^{-1} A_2^t w_2)^{-1} A_2 Q^{-1}$$

or,

$$w_2 A_2 w_1^{-1} = K A_2 Q^{-1} \text{ (say)},$$

which will be seen to have the same meaning as

$$Q^{-1}A_2{}^tKA_2Q^{-1} = Q^{-1}A_2{}^tw_2A_2Q^{-1}.$$

From Equation 9 it is immediately apparent that

$$Q^{-1}A_2{}^tKA_2Q^{-1} = Q^{-1}(Q - Q_1)Q_1{}^{-1} = Q_1{}^{-1} - Q^{-1}$$

and the inverse of the matrix of coefficients of the reduced set of normal equations is given by

$$Q_1{}^{-1} = Q^{-1} + Q^{-1}A_2{}^tKA_2Q^{-1}. \tag{11}$$

Thus, if the number of observations to be omitted from an already completed adjustment is small, it is preferable to calculate the matrix K which would involve the inversion of $(I - A_2Q^{-1}A_2{}^tw_2)$, a $k \times k$ matrix, instead of the very much larger $(n-k) \times (n-k)$ matrix Q_1.

The unbiased estimates of the corrections to the unknowns, viz., $\Delta X'^*$, as a result of dropping k observations, is then

$$\Delta X'^* = -Q_1{}^{-1}F_1. \tag{12}$$

From Equations 12 and 5 we have

$$\begin{aligned} Q\cdot\Delta X^* - Q_1\cdot\Delta X'^* &= -F + F_1 \\ &= -F + F - A_2{}^tw_2f_2 \end{aligned}$$

and

$$Q_1\cdot\Delta X'^* = Q\cdot\Delta X^* + A_2{}^tw_2f_2. \tag{13}$$

Premultiplying $(A_2\cdot\Delta X'^* + f_2)$ by $K^{-1}w_2$ we have

$$\begin{aligned} (I - A_2Q^{-1}A_2{}^tw_2)(A_2\cdot\Delta X'^* + f_2) &= f_2 + A_2\cdot\Delta X'^* - A_2Q^{-1}(Q - Q_1)\Delta Q'^* + \\ &\quad - A_2Q^{-1}(Q_1\cdot\Delta X'^* - Q\cdot\Delta X^*) \\ &= f_2 + A_2\cdot\Delta X^* = v_2. \end{aligned}$$

Therefore,

$$v_2 = f_2 + A_2\Delta X^* \tag{14}$$

or

$$v_2 = (I - A_2Q^{-1}A_2{}^tw_2)(A_2\cdot\Delta X'^* + f_2). \tag{15}$$

Premultiplying Equation 14 by K,

$$Kv_2 = w_2(I - A_2Q^{-1}A_2{}^tw_2)^{-1}(I - A_2Q^{-1}A_2{}^tw_2)(A_2\cdot\Delta X'^* + f_2)$$

and

$$\begin{aligned} Q^{-1}A_2{}^tKv_2 &= Q^{-1}A_2{}^t(A_2\,\Delta^*X' + f_2) \\ &= Q^{-1}(Q - Q_1)\Delta X'^* + Q^{-1}(Q_1\cdot\Delta X'^* - Q\,\Delta X^*) \\ &= \Delta X'^* - \Delta X^* \end{aligned}$$

$$\Delta X'^* = \Delta X^* + Q^{-1}A_2{}^tKv_2. \tag{16}$$

For the n observations,

$$\begin{aligned} v^twv &= (A\cdot\Delta X^* + f)^tw(A\cdot\Delta X^* - f) \\ &= \Delta X^{*t}(A^twA\cdot\Delta X^* + A^twf) - \Delta X^{*t}Q\cdot\Delta X^* + f^twf \end{aligned}$$

$$v^twv = f^twf - \Delta X^{*t}\Delta X^*. \tag{17}$$

For the $(n-k)$ observations,

$$v_1{}^t w_1 v_1 = f_1{}^t w_1 f_1 - \Delta X'^{*t} Q_1 \cdot \Delta X'^{*}. \tag{18}$$

Subtracting Equation 18 from 17 and rearranging slightly,

$$v_1{}^t w_1 v_1 = v^t w v - f_2{}^t w_2 f_2 + \Delta X^{*t} Q \cdot \Delta X^{*} - \Delta X'^{*t} Q_1 \cdot \Delta X'^{*}.$$

Now,

$$\begin{aligned} v_2{}^t K v_2 &= (A_2 \cdot \Delta X^{*} + f_2)^t w_2 (A_2 \cdot \Delta X'^{*} + f_2) \\ &= f_2{}^t w_2 f_2 - \Delta X^{*t} Q \cdot \Delta X^{*} + \Delta X'^{*t} Q_1 \cdot \Delta Q'^{*}. \end{aligned}$$

Therefore,

$$v_1{}^t w_1 v_1 = v^t w v - v_2{}^t K v_2. \tag{19}$$

Finally,

$$v_2 = (I - A_2 Q^{-1} A_2{}^t w_2)(A_2 \cdot \Delta X'^{*} + f_2).$$

That is,

$$\begin{aligned} K v_2 &= w_2 (A_2 \cdot \Delta X'^{*} + f_2) \\ w_2{}^{-1} K v_2 &= A_2 \cdot \Delta X'^{*} + f_2 \end{aligned}$$

or

$$A_2 \cdot \Delta X'^{*} = w_2{}^{-1} K v_2 - f_2.$$

Thus,

$$v' = A_2 \cdot \Delta X'^{*} + f_2 = w_2{}^{-1} K v_2 \tag{20}$$

where v' is the vector of residuals of the omitted observations based on the estimate $\Delta X'^{*}$ of the unknowns ΔX. The matrix K is symmetrical and features in most of the correction terms and, as it is usually of order less than 4×4, it is easier to evaluate than Q_1^{-1}.

It is a simple matter to rewrite the foregoing equations to allow for the addition of further observations to the original set. An example follows to illustrate the application of the correction formulae derived in this paper.

Worked Example

Consider the following trilateration network measured on a photograph for the determination of the plate coordinates of the photo-points 4 and 25. The measured distances, in millimeters, are shown in Table 1 and are shown diagrammatically in Figure 1.

Table 1. Measured Distances in Millimeters

Points A	Points B	Measured Distance	Points A	Points B	Measured Distance
112	114	299.3818	111	113	299.2935
111	112	211.6378	112	113	211.7720
113	114	211.4608	111	114	211.7908
114	4	153.7319	113	4	150.1145
112	4	145.6553	111	4	149.2856
4	25	110.0078	111	25	239.0362
112	25	236.1424	113	25	102.4410
114	25	109.0322			

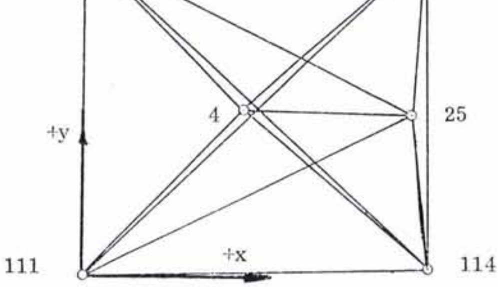

Fig. 1. Trilateration figure used in the example.

$$
A = \begin{bmatrix}
-0.7068 & 0.0000 & 0.0000 & 0.7074 & 0.7068 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.7076 & 0.7067 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0007 & 1.0000 & -0.0007 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & -0.0001 & 1.0000 & 0.0001 & -1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0002 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.7096 & -0.7046 & -0.7096 & 0.7046 & 0.0000 & 0.0000 \\
0.0000 & 0.7266 & 0.6871 & 0.0000 & 0.0000 & -0.7266 & -0.6871 & 0.0000 & 0.0000 \\
0.7091 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.7051 & -0.7091 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.6879 & 0.7258 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 & -0.0065 & 1.0000 & 0.0065 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.8898 & 0.4563 \\
0.4344 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.9008 & -0.4344 \\
0.0000 & -0.0091 & 0.9999 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0091 & -0.9999 \\
0.0000 & 0.0000 & 0.0000 & -0.0084 & -1.0000 & 0.0000 & 0.0000 & 0.0084 & 1.0000
\end{bmatrix}
$$

FIG. 2. The $\boldsymbol{A}$-matrix of the example for use in Equation 1.

The matrices of Equation 1 for this example may be written as in Figure 2 and as follows:

$$\Delta \boldsymbol{X} = (dy_{112}\, dx_{113}\, dy_{113}\, dx_{114}\, dy_{114}\, dx_4\, dy_4\, dx_{25}\, dy_{25})^t$$

$$\boldsymbol{v} = (v_{112-114}\, v_{111-113}\, v_{111-112}\, v_{112-113}\, v_{113-114}\, v_{111-114}\, v_{114-4}\, v_{113-4}\, v_{112-4}\, v_{111-4}\, v_{4-25}\, v_{111-25}\, v_{112-25}\, v_{113-25}\, v_{114-25})^t$$

$$(\boldsymbol{X}) = ((X_{111})\,(Y_{111})\,(X_{112})\,(Y_{112})\,(X_{113})\,(Y_{113})\,(X_{114})\,(Y_{114})\,(X_4)\,(Y_4)\,(X_{25})\,(Y_{25}))^t.$$

that is,

$$(\boldsymbol{X}) = (0.0\ 0.0\ 0.0\ 211.6365\ 211.7694\ 211.4960\ 211.7903\ 0.0347\ 102.6991\ 108.3506\ 212.7039\ 109.0621)^t.$$

$$f = (1.6\ 0.4\ -1.3\ -2.5\ 0.5\ -0.5\ -0.9\ 3.1\ -1.3\ 2.6\ -0.7\ -1.7\ 2.6\ -2.9 -1.0)^t,$$

expressed as μm. The weight matrix $\boldsymbol{w} = \boldsymbol{I}$ and accordingly, the $\boldsymbol{Q}^{-1}$ matrix is shown in Figure 3.

The estimates of the unknowns from Equations 6 and 4 are

$$\boldsymbol{X}^* = (\boldsymbol{X}) + \Delta \boldsymbol{X}^* = (0.0000\ 0.0000\ 0.0000\ 211.6365\ 211.7694\ 211.4960\ 211.7903\ 0.0347\ 102.6991\ 108.3506\ 212.7039\ 109.0621)^t.$$

As $\boldsymbol{X} = (\boldsymbol{X})$, it follows immediately that the vector of corrections for the measured distances $\boldsymbol{v} = \boldsymbol{f}$. The standard deviation of the distance measurement of unit weight is then

$$\sigma_0 = [\boldsymbol{v}^t \boldsymbol{w} \boldsymbol{v}/(n - m)]^{\frac{1}{2}} = 2.88\ \mu\text{m}.$$

If it is now supposed that the distance between points 112 and 25 is doubtful and that we wish to obtain a new estimate of the unknowns from the other fourteen

$$
Q^{-1} = \begin{bmatrix}
0.834 & -0.152 & 0.362 & -0.152 & 0.462 & -0.287 & 0.410 & -0.237 & 0.410 \\
 & 0.766 & -0.491 & -0.106 & -0.381 & 0.152 & -0.031 & 0.085 & -0.379 \\
 & & 1.150 & 0.233 & 0.901 & -0.098 & 0.406 & -0.099 & 0.887 \\
 & & & 0.773 & 0.342 & 0.144 & -0.118 & 0.077 & 0.231 \\
 & & & & 1.244 & -0.192 & 0.457 & -0.149 & 0.935 \\
 & & & & & 0.530 & -0.142 & 0.247 & -0.144 \\
 & Q^{-1}_{ij} = Q^{-1}_{ji},\ i \neq j & & & & & 0.835 & -0.117 & 0.393 \\
 & & & & & & & 0.515 & -0.126 \\
 & & & & & & & & 1.213
\end{bmatrix}
$$

FIG. 3. The $\boldsymbol{Q}^{-1}$ matrix for the example.

$$K = I[I - (0.434\ 0.0\ 0.0\ 0.0\ 0.0\ 0.0\ 0.0\ 0.901\ -0.434)Q^{-1}\begin{bmatrix} 0.434 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.901 \\ -0.434 \end{bmatrix} I]^{-1}$$

$$K = [I - (-0.029\ 0.175\ -0.317\ -0.097\ -0.340\ 0.160\ -0.098\ 0.416\ -0.462)\begin{bmatrix} 0.434 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.901 \\ -0.434 \end{bmatrix}]^{-1}$$

FIG. 4. Evaluation of K for the example.

distances we rearrange A so that A_2 from Equation 7 may be written as

$$A_2 = (0.4344\ 0.0000\ 0.0000\ 0.0000\ 0.0000\ 0.0000\ 0.0000\ 0.9008\ -0.4344).$$

The terms $f_2 = (2.6)$ in μm and $v_2 = (v_{12-25})$. We have here that $w_2 = w_2^{-1} = I$.

The correction procedure now revolves round the evaluation of K as

$$K = w_2(I - A_2Q^{-1}A_2^tw_2)^{-1}$$

evaluated as shown in Figure 4 or

$$K = (1 - 0.563)^{-1} = 2.229.$$

The new estimates of the unknowns may now be calculated as

$$X'^* = X^* + \Delta X^* + Q^{-1}A_2^tKv_2$$

as shown in Figure 5, and the coordinates of the points are shown Table 2.

$$X'^* = \begin{bmatrix} 211.6365 \\ 211.7694 \\ 211.4960 \\ 211.7903 \\ 0.0347 \\ 102.6991 \\ 108.3506 \\ 212.7039 \\ 109.0621 \end{bmatrix} + Q^{-1}\begin{bmatrix} 0.434 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.000 \\ 0.901 \\ -0.434 \end{bmatrix} 2.229 \times 2.6\ \mu m = \begin{bmatrix} 211.6365 \\ 211.7694 \\ 211.4960 \\ 211.7903 \\ 0.0347 \\ 102.6991 \\ 108.3506 \\ 212.7039 \\ 109.0621 \end{bmatrix} + \begin{bmatrix} -0.2 \\ +1.0 \\ -1.8 \\ -0.6 \\ -2.0 \\ +0.9 \\ -0.6 \\ +2.4 \\ -2.7 \end{bmatrix} \mu m$$

FIG. 5. The new estimates of the unknowns.

TABLE 2. POINT COORDINATES FOR X'^*.

Point	*X mm Y*	
111	0.0000	0.0000
112	0.0000	211.6363
113	211.7704	211.4942
114	211.7897	0.0327
4	102.7000	108.3500
25	212.7063	109.0594

$$
Q^{-1}A_2^tKA_2Q^{-1} = \begin{bmatrix}
0.002 & -0.008 & 0.022 & 0.006 & 0.022 & -0.011 & 0.007 & -0.029 & 0.032 \\
 & 0.069 & -0.126 & -0.039 & -0.136 & 0.064 & -0.039 & 0.165 & -0.185 \\
 & & 0.230 & 0.070 & 0.246 & -0.117 & 0.071 & -0.301 & 0.335 \\
 & & & 0.022 & 0.075 & -0.035 & 0.021 & -0.093 & 0.102 \\
 & & & & 0.263 & -0.125 & 0.056 & -0.321 & 0.357 \\
 & & & & & 0.059 & -0.036 & 0.153 & -0.170 \\
 & a_{ij} = a_{ji},\ i \neq j & & & & & 0.022 & -0.093 & 0.103 \\
 & & & & & & & 0.395 & -0.438 \\
 & & & & & & & & 0.486
\end{bmatrix}
$$

$$
Q_1^{-1} = \begin{bmatrix}
0.836 & -0.164 & 0.384 & -0.146 & 0.484 & -0.298 & 0.417 & -0.266 & 0.442 \\
 & 0.835 & -0.617 & -0.145 & -0.517 & 0.216 & -0.070 & 0.250 & -0.564 \\
 & & 1.380 & 0.303 & 1.147 & -0.215 & 0.477 & -0.400 & 1.222 \\
 & & & 0.795 & 0.417 & 0.109 & -0.097 & -0.016 & 0.333 \\
 & & & & 1.507 & -0.317 & 0.533 & -0.470 & 1.292 \\
 & a_{ij} = a_{ji},\ i \neq j & & & & 0.589 & -0.178 & 0.400 & -0.314 \\
 & & & & & & 0.857 & -0.210 & 0.496 \\
 & & & & & & & 0.910 & -0.564 \\
 & & & & & & & & 1.699
\end{bmatrix}
$$

FIG. 6. Evaluation of Equation 11.

Applying Equation 11 we find the correction term and Q_1^{-1} as indicated in Figure 6.

The vector of corrections to the $n-k$ observations used in the solution for X'^* is given by

$$v_1 = f_1 + A_1 \cdot \Delta X'^*.$$

That is,

$$\boldsymbol{v}_1 = (2.5\ \ -0.2\ \ -1.5\ \ -1.5\ \ 0.6\ \ -1.1\ \ -1.0\ \ 2.3\ \ -0.4\ \ 2.8\ \ 0.7\ \ -0.8\ \ -2.0\ \ -1.7)^t,$$

expressed in μm. The correction to the discarded measurement with respect to the revised adjustment is given by Equation 20 as $v' = 5.8\ \mu m$.

Finally, the standard deviation of the distance measurement of unit weight derived from the revised adjustment is calculated from

$$\sigma_0' = [\boldsymbol{v}_1^t \boldsymbol{w}_1 \boldsymbol{v}_1/(n-k-m)]^{\frac{1}{2}} = [(\boldsymbol{v}^t \boldsymbol{w} \boldsymbol{v} - \boldsymbol{v}_2^t \boldsymbol{K} \boldsymbol{v}_2)/(n-k-m)]^{\frac{1}{2}}$$

$$\sigma_0' = 2.63\ \mu m \qquad (n = 15,\ k = 1,\ m = 9).$$

In the example, it will be noted that dx_1, dy_1, and dx_2 have been set equal to zero. This is possible without altering the generality of the solution.

A complete readjustment of the 14 observations of $\boldsymbol{A}_1$ will lead to exactly the same results for $\boldsymbol{Q}_1^{-1}$, $\boldsymbol{X}'^*$, $\boldsymbol{v}_1$ and σ_0' as have been obtained from the application of the correction terms mentioned earlier in this paper.

REFERENCES

Aitken, A. C., (1934). "On least squares and linear combination of observations." *Proc. Roy. Soc. Edinb.*, A, (55): 42–7

Plackett, R. L., (1950). "Some theorems in Least Squares." *Biometrika*, (37): 149–157.