Acta Cryst. (1981). A**37**, 621–624

Least-Squares Refinement of the Best-Plane Parameters

BY TETSUZO ITO

The Institute of Physical and Chemical Research, Wako-shi, Saitama 351, *Japan*

(*Received* 2 *December* 1980; *accepted* 13 *February* 1981)

Abstract

A least-squares procedure to refine the best-plane parameters is described. By linearizing the quadratic constraint imposed on the components of the plane normal, it is shown that the error matrix (variance–covariance matrix) of the best-plane parameters can be obtained in a straightforward way from the coefficients of the normal equation. The error matrix thus obtained is useful for the estimation of the uncertainties in atom-to-plane distances. Numerical examples are given to illustrate the procedure.

Introduction

A least-squares method to determine the best plane through a set of points has been presented by Schomaker, Waser, Marsh & Bergman (1959). Blow (1960) has proposed transforming their equations into an orthonormal metric, and Hamilton (1961) has introduced a non-diagonal weight matrix. However, because of the quadratic constraint imposed on the components of the plane normal, the least-squares problem leads to an eigenvalue equation, which makes a direct evaluation of the error matrix (variance–covariance matrix, or moment matrix) of the parameters difficult. More recently, Scheringer (1971) has shown by introducing the two Eulerian angles of the plane normal that the solution can be given by using standard least-squares routines. On the other hand, Waser, Marsh & Cordes (1973) have derived formulas for the error matrix by applying the standard error-propagation formula to the isotropic positional variances of the atoms to which the plane has been fitted. In the present treatment, the quadratic constraint is expanded into a linear form, which makes it possible to derive the error matrix directly from the coefficients of the normal equation.

Least-squares refinement

Let us consider the plane through n atoms

$$\mathbf{r}^{(i)} = x^{(i)}\mathbf{a} + y^{(i)}\mathbf{b} + z^{(i)}\mathbf{c}, \quad i = 1,2,\ldots,n. \tag{1}$$

If we define the plane by its unit normal

$$\mathbf{m} = m_1\mathbf{a}^* + m_2\mathbf{b}^* + m_3\mathbf{c}^*, \tag{2}$$

and by its distance d_0 from the origin, the atom-to-plane distances are given by

$$d^{(i)} = m_1x^{(i)} + m_2y^{(i)} + m_3z^{(i)} - d_0, \quad i = 1,2,\ldots,n. \tag{3}$$

Then, our least-squares problem is to minimize†

$$S \equiv \tilde{\mathbf{d}}\mathbf{W}\mathbf{d} \tag{4}$$

under the constraint

$$\tilde{\mathbf{m}}\mathbf{G}^*\mathbf{m} = \tilde{\mathbf{n}}\mathbf{H}^*\mathbf{n} = 1, \tag{5}$$

where

$$\tilde{\mathbf{d}} \equiv (d^{(1)} \quad d^{(2)} \ldots d^{(n)}), \tag{6}$$

$$\tilde{\mathbf{n}} \equiv (m_1 \quad m_2 \quad m_3 \quad d_0) \equiv (\tilde{\mathbf{m}} \quad d_0), \tag{7}$$

$$\mathbf{H}^* \equiv \begin{bmatrix} & & & 0 \\ & \mathbf{G}^* & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{8}$$

$\mathbf{G}^*$ is the metric tensor for the reciprocal lattice.

The weight matrix $\mathbf{W}$ in (4) should be chosen as

$$\mathbf{W} = {}^d\mathbf{M}^{-1}, \tag{9}$$

where ${}^d\mathbf{M}$ is the error matrix for the perpendicular distances of the atoms from the plane. Derivation of ${}^d\mathbf{M}$ from the error matrix for the atomic positional parameters has been described in detail by Hamilton (1961). $\mathbf{W}$ is a diagonal matrix unless atom–atom correlations are taken into account.

If we use the constraint (5) as it is, minimization of S leads to an eigenvalue equation (Schomaker, Waser,

† In the following, matrix notation similar to that introduced by Hamilton (1961) will be used. a lower-case bold Univers symbol (*e.g.* **d**) denotes a column matrix, so that its transpose ($\tilde{\mathbf{d}}$) is a row matrix. An upper-case bold Univers symbol (*e.g.* **W**) denotes a square or a rectangular matrix.

0567-7394/81/050621-04$01.00

Marsh & Bergman, 1959; Blow, 1960; Hamilton, 1961; Scheringer, 1971). In the present treatment, (5) is expanded into a linear form around approximate values of the plane parameters $\mathbf{n}^{(0)}$. $\mathbf{n}^{(0)}$ may be obtained from three arbitrary atoms defining the plane or by the 'prevalent incorrect method' discussed in detail by Schomaker *et al.* (1959). We replace $\tilde{\mathbf{n}}$ in (5) by

$$\tilde{\mathbf{n}} = (\tilde{\mathbf{m}}^{(0)} - \tilde{\Delta\mathbf{m}} \quad d_0^{(0)} - \Delta d_0) \equiv \tilde{\mathbf{n}}^{(0)} - \tilde{\Delta\mathbf{n}} \quad (10)$$

and ignore the quadratic terms in $\Delta\mathbf{n}$. Then, (5) reduces to a linear equation

$$\tilde{\mathbf{h}}\Delta\mathbf{n} = 0, \quad (11)$$

where

$$\mathbf{h} \equiv \mathbf{H}^*\mathbf{n}^{(0)}, \quad (12)$$

and $\mathbf{n}^{(0)}$ is assumed to be normalized:

$$\tilde{\mathbf{n}}^{(0)}\mathbf{H}^*\mathbf{n}^{(0)} = \tilde{\mathbf{h}}\mathbf{n}^{(0)} = 1. \quad (13)$$

Since d_0 is independent of the linearization, its approximate value $d_0^{(0)}$ need not be known; it may be put equal to zero.

Now, the function to be minimized may be written

$$F \equiv \tfrac{1}{2}\tilde{\mathbf{d}}\mathbf{W}\mathbf{d} + \lambda\tilde{\mathbf{h}}\Delta\mathbf{n}, \quad (14)$$

where λ is an undetermined multiplier. Minimization of F with respect to $\Delta\mathbf{n}$ gives the normal equation

$$\mathbf{C}\Delta\mathbf{n} = \tilde{\mathbf{Y}}\mathbf{W}\mathbf{d}^{(0)} - \mathbf{h}\lambda, \quad (15)$$

where

$$\mathbf{Y} \equiv \begin{bmatrix} x^{(1)} & y^{(1)} & z^{(1)} & -1 \\ \vdots & \vdots & \vdots & \vdots \\ x^{(n)} & y^{(n)} & z^{(n)} & -1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{\mathbf{y}}^{(1)} \\ \vdots \\ \tilde{\mathbf{y}}^{(n)} \end{bmatrix}, \quad (16)$$

$$\mathbf{d}^{(0)} \equiv \mathbf{Y}\mathbf{n}^{(0)}, \quad (17)$$

$$\mathbf{C} \equiv \tilde{\mathbf{Y}}\mathbf{W}\mathbf{Y}. \quad (18)$$

The normal equation (15) combined with the linearized constraint (11) can easily be solved for λ and $\Delta\mathbf{n}$ to give

$$\lambda = (\tilde{\mathbf{h}}\mathbf{C}^{-1}\mathbf{h})^{-1}\tilde{\mathbf{h}}\mathbf{C}^{-1}\tilde{\mathbf{Y}}\mathbf{W}\mathbf{d}^{(0)} \quad (19a)$$

$$= (\tilde{\mathbf{h}}\mathbf{C}^{-1}\mathbf{h})^{-1}, \quad (19b)$$

$$\Delta\mathbf{n} = \mathbf{B}\mathbf{C}^{-1}\tilde{\mathbf{Y}}\mathbf{W}\mathbf{d}^{(0)} \quad (20a)$$

$$= \mathbf{B}\mathbf{n}^{(0)}, \quad (20b)$$

where

$$\mathbf{B} \equiv \mathbf{I} - \lambda\mathbf{C}^{-1}\mathbf{h}\tilde{\mathbf{h}}. \quad (21)$$

$\mathbf{I}$ is the 4×4 unit matrix. Compared to the standard least-squares solution without constraints, an additional matrix $\mathbf{B}$ has appeared in the solution (20*a*). λ as evaluated by (19) is equal to S in (4) after the least-squares refinement

$$S = \tilde{\mathbf{d}}\mathbf{W}\mathbf{d} = \tilde{\mathbf{n}}\mathbf{C}\mathbf{n} = \lambda\tilde{\mathbf{n}}\mathbf{h} = \lambda. \quad (22)$$

Since the present procedure employs a linear approximation, iteration cycles must be repeated until changes in the best-plane parameters become sufficiently small. It should be noted, however, that iteration is necessary also in Hamilton's (1961) scheme in order to refine the weight matrix.

The matrix $\mathbf{C}$ in (18) turns out to be singular in some special cases; it happens for example when the plane is defined with only three atoms. Since the inverse matrix $\mathbf{C}^{-1}$ is not defined in such cases, the following relations involving $\mathbf{C}^{-1}$ are not valid. We have to eliminate one of the variable parameters by using the constraint relation (11). Then, the problem reduces to a standard linear least-squares problem without constraints.

Estimation of errors

The error matrix (variance–covariance matrix) of the best-plane parameters can be evaluated directly from (20*a*)

$$\begin{aligned} {}^n\mathbf{M} &= \mathbf{B}\mathbf{C}^{-1}\tilde{\mathbf{Y}}\mathbf{W}^d\mathbf{M}\mathbf{W}\mathbf{Y}\mathbf{C}^{-1}\tilde{\mathbf{B}} \\ &= \mathbf{B}\mathbf{C}^{-1}\tilde{\mathbf{B}} = \mathbf{B}\mathbf{C}^{-1}. \end{aligned} \quad (23)$$

We note in (23) that the error matrix $\mathbf{C}^{-1}$ for the ordinary solution without constraints is modified by $\mathbf{B}$ because of the constraint.

It is sometimes necessary to test statistical significance of an atom-to-plane distance $d^{(i)}$ against its standard deviation $\sigma(d^{(i)})$. Since the errors in the best-plane parameters are now known, they should also be taken into account in the estimation of $\sigma(d^{(i)})$ as well as the covariance ${}^d\mathbf{M}_{ii}$ for the atomic positional parameters.

Since $d^{(i)}$ is expressed in matrix notation as

$$d^{(i)} = \tilde{\mathbf{y}}^{(i)}\mathbf{n}, \quad (3')$$

$$\sigma^2(d^{(i)}) = {}^d\mathbf{M}_{ii} + \tilde{\mathbf{y}}^{(i)}{}^n\mathbf{M}\mathbf{y}^{(i)}. \quad (24)$$

The contribution of the second term in (24) is especially important when the atom has a small positional error ${}^d\mathbf{M}_{ii}$, and is away from the center of the atomic group defining the plane.

Waser, Marsh & Cordes (1973) have derived a formula for the standard deviation for the dihedral angle between two best planes. The corresponding formula can easily be derived from the error matrix ${}^n\mathbf{M}$. Since the dihedral angle θ between the two best planes with unit normals $\mathbf{m}$ and $\mathbf{m}'$ is defined by

$$\cos\theta = \tilde{\mathbf{m}}\mathbf{G}^*\mathbf{m}', \quad (25)$$

the standard deviation for θ is given by

$$\sigma^2(\theta) = \frac{1}{\sin^2\theta}(\tilde{\mathbf{m}}'\mathbf{G}^{*m}\mathbf{M}\mathbf{G}^*\mathbf{m}' + \tilde{\mathbf{m}}\mathbf{G}^{*m'}\mathbf{M}\mathbf{G}^*\mathbf{m}), \quad (26)$$

where the 3×3 matrix ${}^m\mathbf{M}$ is the $\mathbf{m}$ part of the 4×4 matrix ${}^n\mathbf{M}$, and correlation between $\mathbf{m}$ and $\mathbf{m}'$ is

ignored. As was pointed out by Waser *et al.* (1973) with respect to their formula, (26) becomes an undetermined 0/0 when θ is zero (*i.e.* when $\mathbf{m} = \mathbf{m}'$). This can be seen from the relation

$$\tilde{\mathbf{m}}\mathbf{G}^{*m}\mathbf{M}\mathbf{G}^{*}\mathbf{m} = \tilde{\mathbf{n}}\mathbf{H}^{*n}\mathbf{M}\mathbf{H}^{*}\mathbf{n} = \tilde{\mathbf{h}}^{n}\mathbf{M}\mathbf{h} = 0, \quad (27)$$

where the difference between $\mathbf{n}$ and $\mathbf{n}^{(0)}$ is ignored. When θ approaches zero, however, we see that the errors in the unit normals $\mathbf{m}$ and $\mathbf{m}'$ directly contribute to the error in θ. From the relations

$$|\Delta\mathbf{m}|^2 = \Delta\tilde{\mathbf{m}}\mathbf{G}^{*}\Delta\mathbf{m}, \quad (28a)$$

$$|\Delta\mathbf{m}'|^2 = \Delta\tilde{\mathbf{m}}'\mathbf{G}^{*}\Delta\mathbf{m}', \quad (28b)$$

we obtain

$$\sigma_0^2(\theta) = \text{Trace}\,(\mathbf{G}^{*m}\mathbf{M}) + \text{Trace}\,(\mathbf{G}^{*m'}\mathbf{M}). \quad (29)$$

In practice, the following criterion will be useful for the choice between the expressions (26) and (29)

$$\sigma(\theta) = \sigma_0(\theta) \quad \text{if } \theta < \sigma_0(\theta). \quad (30)$$

If we are concerned with the angle φ between the plane normal $\mathbf{m}$ and a unit vector defined in the direct lattice

$$\mathbf{l} = l_1\mathbf{a} + l_2\mathbf{b} + l_3\mathbf{c}, \quad (31)$$

the angle and its standard deviation can be calculated from

$$\cos\varphi = \tilde{\mathbf{m}}\mathbf{l}, \quad (32)$$

$$\sigma^2(\varphi) = \frac{1}{\sin^2\varphi}(\tilde{\mathbf{l}}^{m}\mathbf{M}\mathbf{l} + \tilde{\mathbf{m}}^{l}\mathbf{M}\mathbf{m}), \quad (33)$$

where ${}^{l}\mathbf{M}$ is the error matrix for $\mathbf{l}$.

Numerical examples

Computational considerations

A Fortran program was written and applied to several examples. Calculations were performed on a FACOM 230-75 computer of this institute. It has been noted that the symmetry of the ${}^{n}\mathbf{M}$ matrix evaluated as $\mathbf{B}\mathbf{C}^{-1}$ in (23) is a useful check for the precision of computation. In fact, most of the matrix manipulations, especially the inversion of the $\mathbf{C}$ matrix, should be handled in double precision in order to ensure sufficient precision.

If a computer program in the scheme of Hamilton (1961) is already available, it can easily be modified to evaluate the error matrix ${}^{n}\mathbf{M}$ from (21) and (23). Modification of a program in the scheme of Schomaker *et al.* (1959) is a little more difficult, because the matrix $\mathbf{C}$ has not been set up in this scheme.

Example of Hamilton (1961)

Let us first take the two-dimensional example of Hamilton: a problem of fitting a line to four points in the plane. Hamilton has demonstrated the use of a non-diagonal weight matrix in this example. The coordinate axes are orthogonal with unit length, so that

$$\mathbf{H}^{*} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (8)$$

If we start from the same approximate parameters as assumed by Hamilton:

$$\tilde{\mathbf{n}}^{(0)} \equiv (m_1^{(0)} \quad m_2^{(0)} \quad d_0^{(0)}) = (0{\cdot}1961 \quad -0{\cdot}9806 \quad 0{\cdot}0),$$

the $\mathbf{C}$ matrix (18) is already given in the paper.

Then

$$\mathbf{C}^{-1} = \begin{bmatrix} 0{\cdot}14991 & -0{\cdot}51801 & -0{\cdot}22244 \\ -0{\cdot}51801 & 2{\cdot}67207 & 1{\cdot}21114 \\ -0{\cdot}22244 & 1{\cdot}21114 & 0{\cdot}55842 \end{bmatrix} \times 10^{-1},$$

$$\tilde{\mathbf{h}} = (0{\cdot}1961 \quad -0{\cdot}9806 \quad 0{\cdot}0), \quad (12)$$

$$\lambda = 3{\cdot}60441, \quad (19)$$

$$\mathbf{B} = \begin{bmatrix} 0{\cdot}96202 & 0{\cdot}18993 & 0{\cdot}0 \\ 0{\cdot}19238 & 0{\cdot}03798 & 0{\cdot}0 \\ 0{\cdot}08703 & -0{\cdot}43519 & 1{\cdot}0 \end{bmatrix}, \quad (21)$$

$$\Delta\tilde{\mathbf{n}} = (0{\cdot}0024 \quad 0{\cdot}0005 \quad 0{\cdot}4438), \quad (20b)$$

$$\tilde{\mathbf{n}} = (0{\cdot}1937 \quad -0{\cdot}9811 \quad -0{\cdot}4438). \quad (10)$$

The solution $\mathbf{n}$ is in agreement with that of Hamilton within ± 1 in the last digit; one cycle of refinement was sufficient in this example. The effect of the constraint on the least-squares solution is apparent on the matrix $\mathbf{B}$, which is significantly deformed from the unit matrix. The error matrix of the parameters is

$${}^{n}\mathbf{M} = \begin{bmatrix} 0{\cdot}4583 & 0{\cdot}0917 & 0{\cdot}1604 \\ 0{\cdot}0917 & 0{\cdot}0183 & 0{\cdot}0321 \\ 0{\cdot}1604 & 0{\cdot}0321 & 0{\cdot}1199 \end{bmatrix} \times 10^{-2}, \quad (23)$$

from which the standard deviations of the parameters are

$$(\sigma(m_1) \quad \sigma(m_2) \quad \sigma(d_0)) = (0{\cdot}0677 \quad 0{\cdot}0135 \quad 0{\cdot}0346),$$

and correlation between m_1 and m_2 is perfect as expected,

$${}^{n}\mathbf{M}_{12}/\sigma(m_1)\sigma(m_2) = 1{\cdot}00.$$

Example of Waser et al. (1973)

The example of Waser *et al.* was next taken up in order to check the agreement between the results obtained by the two methods. The unit-cell dimensions,

the atomic coordinates, and their isotropic standard deviations were taken from the paper of Waser *et al.* Calculations were performed for the same two planes as used by them. In this case, the weight matrix **W** is diagonal with $\mathbf{W}_{ii} = 1/(\sigma^{(i)})^2$, where $\sigma^{(i)}$ is the isotropic standard deviation. The starting values of the best-plane parameters were obtained by the 'incorrect method' with unit weights. Iteration converged very quickly for both planes; the parameters changed about the order of the standard deviations in the first cycle, but the shifts in the second cycle were almost zero (less than 0·3% of the standard deviations).

The best-plane parameters with their standard deviations are given in Table 1, and the error matrix for the parameters is given in Table 2. The standard deviations for the origin-to-plane distances, 0·0032 and 0·0028 Å for planes 1 and 2 respectively, are in exact agreement with those obtained by Waser *et al.* A direct comparison of the other results is difficult, because different coordinate axes were used in the two treatments. The atom-to-plane distances with their standard deviations estimated from (24) are given in Table 3. It can be seen from the table that the standard deviations for the atoms away from the plane are larger than those for the atoms defining the plane.

The dihedral angle θ between planes 1 and 2, and its standard deviation $\sigma(\theta)$ were evaluated from (25) and (26) to give $\theta = 2{\cdot}070$ and $\sigma(\theta) = 0{\cdot}074°$. They are in good agreement with the results, $\theta = 2{\cdot}065$ and $\sigma(\theta) = 0{\cdot}074°$, reported by Waser *et al.* The small difference in θ can be ascribed to the slightly different weighting schemes adopted in the two treatments. Waser *et al.* have reported also the value $\sigma(\theta) = 0{\cdot}085°$ which was obtained by including the effect of the atoms shared by the two planes. However, the effect may be neglected for most purposes where the standard deviation multiplied by a factor of two to three is used for comparison.

Table 1. *Best-plane parameters with their standard deviations in parentheses* (Å)

	Plane 1	Plane 2
m_1	2·366 (9)	2·688 (8)
m_2	7·345 (7)	7·277 (7)
m_3	7·173 (6)	6·960 (6)
d_0	10·150 (3)	10·118 (3)

Table 2. *Error matrix (variance–covariance matrix) for the best-plane parameters* ($\times 10^8$) *in* Å^2

Plane 1

	m_1	m_2	m_3	d_0
m_1	7555	−1912	−4675	−217
m_2	−1912	5521	−1211	2248
m_3	−4675	−1211	4031	−909
d_0	−217	2248	−909	1024

Plane 2

	m_1	m_2	m_3	d_0
m_1	5740	−1778	−3685	−1517
m_2	−1778	4629	−819	1601
m_3	−3685	−819	3308	431
d_0	−1517	1601	431	782

Table 3. *Atom-to-plane distances with their standard deviations in parentheses* (Å)

The atoms defining the plane are marked with an asterisk.

Plane 1		Plane 2	
C(4)*	0·003 (2)	C(4)*	0·007 (2)
C(9)*	0·005 (2)	C(9)*	−0·014 (2)
N(1)*	−0·007 (2)	C(8)*	0·013 (3)
C(2)*	0·013 (2)	C(7)*	0·001 (3)
C(3)*	−0·009 (2)	C(6)*	−0·010 (3)
C(8)	0·081 (3)	C(5)*	0·004 (2)
C(7)	0·096 (4)	N(1)	0·012 (3)
C(6)	0·062 (4)	C(2)	0·077 (3)
C(5)	0·026 (3)	C(3)	0·047 (3)

References

BLOW, D. M. (1960). *Acta Cryst.* **13**, 168.

HAMILTON, W. C. (1961). *Acta Cryst.* **14**, 185–189.

SCHERINGER, C. (1971). *Acta Cryst.* B**27**, 1470–1472.

SCHOMAKER, V., WASER, J., MARSH, R. E. & BERGMAN, G. (1959). *Acta Cryst.* **12**, 600–604.

WASER, J., MARSH, R. E. & CORDES, A. W. (1973). *Acta Cryst.* B**29**, 2703–2708.