JAN 21 1963

LADC 5581

A NOTE ON LEAST SQUARES ESTIMATION OF PARAMETERS WHEN

FITTING A PLANE*

LADC-5581

By Aaron Goldman

University of California, Los Alamos Scientific Laboratory
Los Alamos, New Mexico, U. S. A.

Facsimile Price \$	1.10
Microfilm Price \$	.80
Available from the Office of Technical Services Department of Commerce Washington 25, D. C.	

Some important articles by Schomaker, Waser, Marsh, and Bergman

**

** Henceforth abbreviated by S.M.B.

(1959) and Hamilton (1961) provide a method for fitting a plane through a set of data points by the method of least squares. Their mathematically sound technique has the distinct disadvantage of being difficult to follow because of the necessity of rather advanced mathematics, particularly in the methods of eigenvalue solution. It is the purpose of this note to describe a less involved method and an IBM 7090 code that is available to solve the problem.

Matrix notation will be used because of the rather cumbersome summations that arise.

The problem is to find the best values of the parameters in the model

$$\begin{matrix} X & A & = & E \\ n \times r & r \times 1 & & n \times 1 \end{matrix}$$

where the subscripts under the terms denote the dimensions of the matrix,

* Work performed under the auspices of the U. S. Atomic Energy Commission.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:
A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information disclosed herein may not infringe upon privately owned rights;
B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.
As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor, prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

This paper was submitted for publication in the open literature at least 6 months prior to the issuance date of this Microcard. Since the U.S.A.E.C. has no evidence that it has been published, the paper is being distributed in Microcard form as a preprint.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

$$X = \begin{bmatrix} x_{11} & \dots & x_{k1} & 1 \\ \vdots & & \vdots & \vdots \\ x_{1n} & \dots & x_{kn} & 1 \end{bmatrix}, \quad A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \\ a_{k+1} \end{bmatrix}, \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

x_{ij} is an observed random variable,

a_i is a parameter,

e_j is a random error variable with mean 0,

$i = 1, \dots, k, k + 1,$

$r = k + 1,$

and

$j = 1, 2, \dots, n.$

Thus there are n observations with $k + 1$ parameters to be estimated.

It happens that in many crystallographic problems the variables are correlated; consequently the symmetric variance covariance matrix of x_{ij} is

$$V = \begin{matrix} & \begin{bmatrix} V_{11} & V_{12} & \dots & V_{1n} \\ & V_{22} & \dots & V_{2n} \\ & & & V_{nn} \end{bmatrix} \\ \begin{matrix} V = \\ kn \times kn \end{matrix} & \end{matrix}.$$

The p th row of V represents the covariances between x_{pt} and the remaining variables. The subscript p runs from 1 through k in each subdivision, identified as V_{ij} , and the subscript t runs from 1 through n . For example

$$V_{11} = \begin{bmatrix} \sigma_{x_{11}x_{11}} & \sigma_{x_{11}x_{21}} & \dots & \sigma_{x_{11}x_{k1}} \\ \sigma_{x_{21}x_{11}} & \sigma_{x_{21}x_{21}} & \dots & \sigma_{x_{21}x_{k1}} \\ \vdots & \vdots & \dots & \vdots \\ \sigma_{x_{k1}x_{11}} & \sigma_{x_{k1}x_{21}} & \dots & \sigma_{x_{k1}x_{k1}} \end{bmatrix}$$

where $\sigma_{x_{ab}x_{cd}}$ identifies the covariance between x_{ab} and x_{cd} .

Let

$$A^* = \begin{bmatrix} A_1 & \varphi & \varphi & \dots & \varphi \\ \varphi & A_1 & \varphi & \dots & \varphi \\ \vdots & \vdots & \vdots & & \vdots \\ \varphi & \varphi & \varphi & \dots & A_1 \end{bmatrix}$$

$n \times nk$

where

$$A_1 = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \dots & \hat{a}_k \end{bmatrix},$$

$1 \times k$

$\hat{a}_1$ is an estimate of a_1 , and φ is the null vector.

By propagation of error and assuming the elements in A are constant, then the new variance covariance matrix of the quantity $\sum_{j=1}^k a_j x_{1j} + a_{k+1}$, where $j = 1, 2, \dots, n$, is given as

$$W = A^* V A$$

$n \times n$

where A^* denotes the transpose of A .

It is apparent from the original model that a straightforward solution by least squares leads to a set of homogeneous equations. The solution may be found by setting

$$\hat{a}_k = \sqrt{1 - \sum_{i=1}^{k-1} \hat{a}_i^2}.$$

The normal equations to be solved are given by

$$X^T W^{-1} X \hat{A} = \phi$$

where

$$\hat{A} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_k \\ \hat{a}_{k+1} \end{bmatrix}$$

and W^{-1} is the inverse matrix of W .

The technique given by SWMB and Hamilton provide a solution by solving for eigenvalues. An alternative method is now presented.

Instead of solving the nonlinear (nonlinear in $\hat{a}$) equations, the original model is re-written as a Taylor series expansion about the parameters.

Let

$$y_i = \sum_{j=1}^{k-1} a_j x_j + x_k \sqrt{1 - \sum_{j=1}^{k-1} a_j^2} + a_{k+1},$$

$\hat{a}_{p,0}$ denote the initial guess of the p th parameter, and $\Delta y_{i,0}$ denote

the resulting value of the function at the i th data point.

Then the model may be written in the form of the expansion

$$\Delta y_{i,0} = \sum_{p=1}^k \left(\frac{\partial y}{\partial a_p} \bigg|_{i,0} \right) \Delta a_{p,1}$$

where $\left(\frac{\partial y}{\partial a_p} \bigg|_{i,0} \right)$ denotes the partial derivation of y with respect to the p th parameter evaluated at $(\hat{a}_{1,0}, \hat{a}_{2,0}, \dots, \hat{a}_{k,0})$ for the i th data point. The values of $\Delta a_{p,1}$ are to be estimated by least squares and new estimates of a_p are found by setting

$$\hat{a}_{k,q} = \hat{a}_{k,q-1} + \Delta a_{k,q-1}$$

at the q th iteration.

Reverting to matrix notation, let

$$\underset{n \times 1}{Y} = \begin{bmatrix} \Delta y_{1p} \\ \Delta y_{2p} \\ \vdots \\ \Delta y_{np} \end{bmatrix};$$

$$\underset{k \times n}{B'} = \begin{bmatrix} \frac{\partial y}{\partial a_1} \bigg|_{i=1,0} & \frac{\partial y}{\partial a_1} \bigg|_{i=2,0} & \dots & \frac{\partial y}{\partial a_1} \bigg|_{i=n,0} \\ \frac{\partial y}{\partial a_2} \bigg|_{i=1,0} & \frac{\partial y}{\partial a_2} \bigg|_{i=2,0} & \dots & \frac{\partial y}{\partial a_2} \bigg|_{i=n,0} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y}{\partial a_{k-1}} \bigg|_{i=1,0} & \frac{\partial y}{\partial a_{k-1}} \bigg|_{i=2,0} & \dots & \frac{\partial y}{\partial a_{k-1}} \bigg|_{i=n,0} \\ 1 & 1 & \dots & 1 \end{bmatrix};$$

and

$$Q = \begin{bmatrix} \Delta a_1 \\ \Delta a_2 \\ \vdots \\ \Delta a_{k-1} \\ \Delta a_{k+1} \end{bmatrix} \cdot$$

The linear normal equations that are solved on the first iteration are given as

$$B' W^{-1} B Q = B' W^{-1} Y$$

and the first solution for Q is

$$Q = (B' W^{-1} B)^{-1} B' W^{-1} Y.$$

Successive iterations are made until the elements in Q become very small (e.g. 10^{-7}).

The variance and covariances of $\hat{a}_i$ are obtained directly from the elements of $(B' W^{-1} B)^{-1}$. The variance of $\hat{a}_k$ is obtained by propagation of error. If the assumption is made that the x_{ij} are independent and have a constant unknown variance, then the variance and covariances of $\hat{a}_i$ are obtained by multiplying the corresponding element of the B'B matrix times s^2 where

$$s^2 = \frac{\sum_{i=1}^n \Delta y_{if}^2}{n - k}$$

and f is the final iteration. The code gives the user the option of either assumption. Examples given by SWMB and Hamilton were used in verifying the code.

Similar results were obtained.

In order to summarize the technique, the following example of an experiment conducted by Howard Cady at the Los Alamos Scientific Laboratory is presented.

The plane which was fitted represents the nitramine group in the structure of the alpha polymorph of HMX. The complete detailed results will appear in a future issue of ACTA CRYSTALLOGRAPHICA along with an unweighted planar fit of the data presented.

The variance-covariance matrix is given as follows:

	x_{11}	x_{21}	x_{31}	x_{12}	x_{22}	x_{32}	x_{13}	x_{23}	x_{33}	x_{14}	x_{24}	x_{34}
x_{11}	2.9(-7)	7.0(-10)	0.0	-9.3(-8)	1.1(-8)	-3.7(-9)	2.9(-8)	-1.6(-8)	-4.0(-8)	3.4(-8)	-9.1(-9)	1.3(-8)
x_{21}		6.8(-8)	0.0	3.9(-9)	-2.6(-9)	9.0(-9)	7.3(-9)	-3.9(-9)	8.0(-8)	-9.5(-9)	-5.0(-9)	6.6(-8)
x_{31}			0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
x_{12}				3.8(-7)	2.6(-9)	8.7(-8)	3.1(-8)	7.5(-9)	-2.5(-8)	7.1(-8)	2.7(-8)	7.3(-8)
x_{22}					9.0(-8)	2.2(-8)	1.5(-8)	-3.4(-9)	6.0(-8)	4.6(-9)	-9.6(-9)	3.9(-8)
x_{32}						3.0(-6)	-3.5(-8)	-2.7(-8)	1.2(-6)	1.7(-7)	-4.6(-9)	1.2(-6)
x_{13}							2.5(-7)	-8.4(-9)	2.8(-8)	6.6(-9)	2.0(-8)	5.9(-8)
x_{23}								7.8(-8)	3.8(-9)	-5.5(-9)	-4.8(-9)	-1.6(-8)
x_{33}									3.8(-6)	-1.6(-7)	-3.9(-9)	1.9(-6)
x_{14}										2.8(-7)	-9.6(-9)	-4.1(-8)
x_{24}											1.1(-7)	-1.9(-8)
x_{34}												3.7(-6)

The set of data points include

$$\begin{aligned}
 x_{11} &= 1.808(-2) & x_{12} &= 2.10(-2) & x_{13} &= 4.938(-2) & x_{14} &= 9.035(-21) \\
 x_{21} &= -5.989(-2) & x_{22} &= -7.264(-2) & x_{23} &= -8.774(-2) & x_{24} &= -6.701(-2) \\
 x_{31} &= 9.874(-2) & x_{32} &= 3.2172(-1) & x_{33} &= 4.091(-1) & x_{34} &= 4.2248(-1)
 \end{aligned}$$

A three-dimensional plane was fitted that gave the following results:

Initial guess:

$$a_1 = 0$$

$$a_2 = 0$$

$$a_4 = 0$$

Final results:

Evaluation of fitted plane:

$\hat{a}_1 = - .1511$	$s_{\hat{a}_1} = .0033$	(11,21,31)	1.412 (-4)
$\hat{a}_2 = .9869$	$s_{\hat{a}_2} = .0005$	(12,22,32)	-9.430 (-4)
$\hat{a}_3 = .5576$	$s_{\hat{a}_3} = .0011$	(13,23,33)	3.035 (-4)
$\hat{a}_4 = .5646$	$s_{\hat{a}_4} = .0004$	(14,24,34)	3.995 (-4)

where $s_{\hat{a}_i}$ is the standard deviation of the estimate of a_i , and the triple a_i identifies the data point.

Copies of the code may be obtained by writing to the author at
Box 1663, Los Alamos, New Mexico.

Acknowledgments are made to Dr. R. K. Zeigler and Dr. R. H. Moore
at the Los Alamos Scientific Laboratory for their assistance in the solu-
tion of this problem.

REFERENCES

CADY, H., LARSON A., Cromer, D. To be published in Acta Cryst.

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

MOORE, R. H., ZEIGLER, R. K. (1959). LA-2367 The Solution of the General
Least Squares Problem with Special Reference to High-Speed Computers.

SCHOMAKER, V., WASER, J., MARSH, R. E. and BERGMAN, G. (1959). Acta Cryst.
12, 600.