

17.X-8 PRECISION AND ACCURACY IN REFINEMENT BY THE RIETVELD METHOD. By E. Prince, National Bureau of Standards, Washington, DC 20234, U. S. A.

The method of least squares is based on the assumption that an observation may be represented by  $y_i = M_i(x) + e_i$ , where  $M_i$  represents a model that is a function of a vector,  $x$ , of parameters that are constants of nature whose values are unknown, and  $e_i$  is an error drawn at random from a population with zero mean. If the linear function  $M_i(x) = M_i(x') + \sum_j A_{ij}(x_j - x'_j)$ , where  $A_{ij} = \partial M_i / \partial x_j$ , is a good approximation to  $M_i(x)$  for some  $x'$ , sufficiently close to the true  $x$ , and if the variance-covariance matrix of the observations is  $V_y$ , then the Gauss-Markov theorem states that  $(\hat{x} - x') = H[y - M(x')]$ , where  $H = (A^T W A)^{-1} A^T W$  and  $W = V_y^{-1}$ , is an unbiased estimate of  $x$  for which the variances of all elements of  $x$  are minimized. The variance-covariance matrix for  $x$  is  $H V_y H^T = (A^T W A)^{-1}$ .

The zero means of the populations from which the  $e_i$  are drawn imply that the model is correct, and, if it is correct, the quantity  $[y - M(\hat{x})]^T W [y - M(\hat{x})] / (\underline{n} - \underline{p})$ , where  $\underline{n}$  is the total number of data points and  $\underline{p}$  is the number of parameters, will be close to one. Higher values are indications of lack of fit due to systematic error or an inadequate model.

In the procedure for structure refinement from powder data that was introduced by Rietveld, [J. Appl. Cryst. 2, 65-71 (1969)] the model is  $M_i(x) = b_i + \sum_k I_k G_{ik}$ , where  $b_i$  is background,  $I_k$  is the integrated intensity of a Bragg reflection, and  $G$  is a peak shape function, normalized so that its sum over a Bragg peak is equal to one. The sum is taken over all peaks that can contribute to point  $i$ . In this case  $W$  is a diagonal matrix, with  $W_{ii} = 1/y_i$ , because the value of any  $e_i$  is unaffected by the value of any other. Many workers have observed [e. g. Baharie & Pawley, J. Appl. Cryst. 16, 404-406 (1983)] that some parameters, especially thermal parameters, are not plausibly accurate within the precision indicated by the estimated standard deviation computed by the standard procedure. For this reason there have been a number of attempts to define a number greater than one by which the estimated standard deviation can be multiplied to give a reasonable indication of accuracy. [Pawley, J. Appl. Cryst. 13, 630-633 (1980); Scott, J. Appl. Cryst. 16, 159-163 (1983); Cooper, Z. Kristallogr. 164, 157-158 (1983)]. The problem with this approach is that the data contain no information concerning the correlations between the estimated parameters and the systematic errors that have not been allowed for in the model, so that potential biases cannot be evaluated.

Confusion has arisen in recent years because the vector of residuals,  $y - M(\hat{x})$ , is numerically identical to the vector of observations,  $y - M(x')$ , when the model has been refined to convergence. Designating by  $P$  the matrix  $A(A^T W A)^{-1} A^T W$ , the residuals have a variance-covariance matrix  $V_r = (I - P)V_y(I - P)^T$ , which is not diagonal, so that the errors in the residuals are correlated, with particularly strong correlations within a single Bragg peak. The confusion between observations and residuals has led some workers to suggest use of a non-diagonal weight matrix and others to use a two stage procedure in which integrated intensities are estimated from the pattern and then parameters are estimated from integrated intensities. [Albinati & Willis, J. Appl. Cryst. 15, 361-374 (1982); Cooper, Acta. Cryst. A38, 264-269 (1982)] Actually,  $W_{ii} = 1/y_i$  is exactly the weight matrix prescribed by the Gauss-Markov theorem, and the two stage procedure can be shown to be mathematically identical to the Rietveld procedure provided the weight matrix for the second stage properly accounts for the correlated errors in the integrated intensities of overlapping peaks.

17.X-9 ENTROPY MAXIMIZATION: AN ALTERNATIVE TO SQUARES MINIMIZATION. By D.M. Collins, Department of Chemistry, Texas A&M University, College Station, Texas 77843, USA.

The method of least squares has been a powerful means of fitting a model to a set of observations. Its success notwithstanding, least squares suffers a number of deficiencies which invite alternative methods.

Shannon's entropy measures the ignorance of an observer concerning the outcomes of a system with possible states of known probability. Jaynes has extended the entropy formulation to include prior distributions. With a fixed set of experimental observations as a (normalized) prior distribution  $m$ , and calculated values taken to be a (normalized) distribution  $p$ ,

$$H = -\sum p_i \ln[p_i/m_i]$$

is a negative measure of the information in the difference between the two sets.

If for an unspecified positive-definite observation  $I_i$  the corresponding calculated value is  $J_i$ , the entropy

$$S = -\sum \frac{I_i}{(\sum I_i)} \ln \frac{I_i/\sum I_i}{J_i/\sum J_i}$$

is to be maximized with respect to the parameters of the model. This will be illustrated by application to crystallographic problems of two major types, including the case in which it is necessary to express  $I$  in a Taylor series expansion.

This work has been supported in part by the Robert A. Welch Foundation through grant A-742.

17.X-10 PRECISE AND ACCURATE ESTIMATION OF CRYSTALLOGRAPHIC PARAMETERS BY MAXIMUM LIKELIHOOD AND MINMAX METHODS BY G.B. Mitra, (Mrs) Rabea Ahmed and Prabal Das Gupta, Indian Association for the Cultivation of Science, Calcutta - 700 032, INDIA.

A combination of Maximum likelihood and MinMax methods has been used in determining extremely precise and accurate values of the following crystallographic parameters :- (1) Lattice constants of HgClBr and a  $\beta$ -lactam derivative (2) Debye Temperature factors of HgClBr and the  $\beta$ -lactam derivative (3) Determination of particle size and strain in pure lead quenched from melt with intermediate rapidity from variance and fourth moment of line profiles against range plots. The method consists of testing the hypothesis of zero systematic error in experimental points chosen and then of drawing a suitable minmax line through them for obtaining constant bounded errors in the final measurements made. The extent of uncorrected systematic error at experimentally observed points and the overall error of measurement were determined for each case. The source for uncorrected systematic error in all the cases seemed to be the unknown contributions to the diffused scattering of the background. The results have been reviewed also in the light of (1) the method of maximum correlation (2) the method of moments and (3) the methods of minimum chisquared ( $\chi^2$ ) tests. It is concluded that a suitable combination of maximum likelihood method and the minmax method is the best possible method. An extension of this technique to crystal structure refinement has been suggested.