VSRT MEMO #042

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
HAYSTACK OBSERVATORY
WESTFORD, MASSACHUSETTS 01886

June 25, 2008

Telephone: 781-981-5407
Fax: 781-981-0590

To: VSRT Group
From: Alan E.E. Rogers
Subject: Least squares fitting of a curve to the ozone spectrum

If we assume that a theoretical spectrum of a fixed shape can adequately represent the data when adjusted in amplitude and offset the problem is mathematically similar to finding the slope and intercept of a straight line fit to data. The algebra is most compact suing the following matrix notation:

y is a column vector representing the spectrum

x is a column vector with 2 elements, the first representing the amplitude and the second the offset.

A is a "design" matrix

ε is colun vector of errors

so that

$$y = Ax + \varepsilon$$

The least squares solution is the result of minimizing

$$Q = \varepsilon^T \varepsilon = (y - Ax)^T (y - Ax) = y^T y - x^T A^T y - y^T Ax + x^T A^T Ax$$

setting the derivative of Q with respect to x to zero

$$A^T Ax = A^T y$$

and we obtain the estimate, $\hat{x}$,

$$\hat{x} = \left(A^T A\right)^{-1} A^T y$$

an error estimate is obtained from

$$\left\langle (\hat{x} - x)(\hat{x} - x)^T \right\rangle^{1/2} = \left(A^T A\right)^{-1} A^T \varepsilon\varepsilon^T A \left(A^T A\right)^{-1} = \left[\left(A^T A\right)^{-1}\right]^{1/2} \sigma$$

when the errors are uncorrelated and have a standard deviation of σ. The $\langle\ \rangle$ brackets denote a statistical average.

If the theoretic spectrum is z_i and the measured spectrum y_i the amplitude estimate is

$$\left(a_4 a_0 - a_1 a_3\right)/d$$

where $a_4 = \sum_{i=0}^{N-1} z_i y_i$

$a_0 = N$

$a_1 = \sum z_i$

$a_3 = \sum y_i$

$d = a_0 a_2 - a_1^2$

The error estimate in amplitude

$d^{-1/2}\sigma$

Reference:

Bevington, P.R., Robinson, D.K., “Data reduction and error analysis the physical sciences,” McGraw-Hill, 1992.