EISCAT
TECHNICAL
NOTES

LEAST MEAN SQUARE FITTING OF DATA TO PHYSICAL MODELS

T. Hagfors

KIRUNA
Sweden

EISCAT TECHNICAL NOTE No. 2
JANUARY 1978

LEAST MEAN SQUARE FITTING OF DATA TO PHYSICAL MODELS

T. Hagfors

Summary.
The note describes a general least mean square fitting procedure for adjusting parameters in a physical model to make the model predictious best fit the results of an observation. The method should find application to such EISCAT problems as the fitting of theoretical spectra (or correlation functions) to the corresponding observations or to the deduction of antenna pointing deficiencies from radio astronomy observations.

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/187 40
TELEX 8754 GEOFYSK S

LEAST MEAN SQUARE FITTING OF DATA TO PHYSICAL MODELS

Suppose we desire to determine an approximation to a set of parameters $\{x_1 \ldots . x_N\} = \vec{x}$ from a set of observational data $\{A_1 \ldots . A_M\} = \vec{A}$. The physical model predicts that the outcome of the observations, given $\vec{x}$ should be:

$$\vec{B} = \vec{B}(\vec{x}) = \{B_1(\vec{x}), B_2(\vec{x}) \ldots . . B_M(\vec{x})\} \tag{1}$$

If the observation can be made with infinite accuracy clearly:

$$\vec{B}(x) - \vec{A} = 0 \tag{2}$$

for the correct value of the parameters $\vec{x}$.

Since the observed values are generally noisy there will not be any value of $\vec{x}$ which can make $\vec{B}(x) - \vec{A} = 0$ unless N = M. We, therefore, imagine that we make an estimate of $\vec{x}$, denoted by $\vec{y}$, which will make the difference:

$$\sum_{i=1}^{M} W_i (A_i - B_i(\vec{y}))^2 = \min \tag{3}$$

where W_i are weight factors to be determined.

We next imagine that the best estimate is determined by means of an iteration procedure with increasing accuracy and that the approximation are:

$$\vec{y}_0 , \vec{y}_1 , \ldots .$$

The "correction" from one approximation to the next will be denoted by $\delta\vec{y}_k$ such that:

$$\vec{y}_{k+1} = \vec{y}_k + \delta\vec{y}_k \tag{4}$$

We can then write:

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/187 40
TELEX 8764 GEOFYSK S

$$\vec{B}(\vec{y}_{k+1}) = \vec{B}(\vec{y}_k) + (\delta\vec{y}_k \cdot \nabla)\vec{B}(\vec{y})\Big|_{\vec{y}=\vec{y}_k} \tag{5}$$

The parameter correction $\delta\vec{y}_k$ is determined from the N equations:

$$\sum_{i=1}^{M} W_i (A_i - B_i(\vec{y}_k) - (\delta\vec{y}_k \cdot \nabla) B_i(\vec{y})\Big|_{\vec{y}=\vec{y}_k}) \frac{\partial}{\partial y_e} B_i(\vec{y})\Big|_{\vec{y}=\vec{y}_k} = 0 \tag{6}$$

Suppressing the iteration index we obtain the following equations for $\delta\vec{y}$:

$$\sum_{i=1}^{M} W_i (A_i - B_i(\vec{y}_k)) \frac{\partial B_i(\vec{y}_k)}{\partial y_e} - \sum_{i=1}^{M} (\delta\vec{y} \cdot \nabla) B_i(\vec{y}_k) \frac{\partial}{\partial y_e} B_i(\vec{y}_k) \cdot W_i \equiv 0 \tag{7}$$

We now introduce the following tensorial notation:

$$T_{\ell m} = \sum_{i=1}^{M} \left(\frac{\partial B_i(\vec{y}_k)}{\partial y_m} \quad \frac{\partial B_i(\vec{y}_k)}{\partial y_e} W_i\right) \tag{8}$$

and the following vectorial notation:

$$V_\ell = \sum_{i=1}^{M} (A_i - B_i(\vec{y}_k)) \frac{\partial B_i(\vec{y}_k)}{\partial y_e} W_i \tag{9}$$

For the general non-linear case both $T_{\ell m}$ and V_ℓ are functions of iteration number. When $\vec{B}(\vec{y})$ is a linear function of $\vec{y}$ $T_{\ell m}$ and V_ℓ do <u>not</u> depend on iteration number. Returning to the general case we have:

$$\vec{V}(\vec{y}_k) = \tilde{T}\delta\vec{y}_k \tag{10}$$

This can be inverted to give:

$$\delta\vec{y}_k = T^{-1}(\vec{y}_k) \cdot \vec{V}(\vec{y}_k) \tag{11}$$

and the next approximation to $\vec{y}$ hence becomes:

$$\vec{y}_{k+1} = \vec{y}_k + T^{-1}(\vec{y}_k) \cdot \vec{V}(\vec{y}_k) \tag{12}$$

For the particular case when $\vec{B}(\vec{x})$ is a linear function of $\vec{x}$

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/187 40
TELEX 8754 GEOFYSK S

no iteration is required and the final value of $\vec{y}$ is obtained from

$$\vec{y} = T^{-1}\vec{V} \tag{13}$$

where

$$V_\ell = \sum_{i=1}^{M} (A_i - B_{i0}) \cdot \frac{\partial B_i}{\partial y_\ell} \cdot W_i \tag{14}$$

where B_{i0} is the value of B_i which is obtained by putting $\vec{y} = 0$. The elements of the matrix T are derived from the coefficients of the linear elements.

In the linear case where an explicit form of $\vec{y}$ can be obtained it is also possible to estimate the uncertainty in the determined values. We estimate the root mean square deviation of $\vec{y}$ from the ensemble average as follows:

The deviation from the mean of the deduced values is caused by random deviations in the observed values from their means:

$$\Delta y_j = \sum_{\ell=1}^{N} (T^{-1})_{j\ell} \Delta V_\ell$$

$$= \sum_{\ell=1}^{N} (T^{-1})_{j\ell} \sum_{i=1}^{M} \Delta A_i \frac{\partial B_i}{\partial y_\ell} W_i \tag{15}$$

The terms of the covariance matrix of the deviations of the estimates from their means become:

$$\langle \Delta y_j \Delta y_k \rangle = \sum_{\ell=1}^{N} \sum_{m=1}^{N} (T^{-1})_{j\ell} (T^{-1})_{km} \sum_{i=1}^{M} \sum_{n=1}^{M} \langle \Delta A_i \Delta A_n \rangle \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_n}{\partial y_m} W_i W_n \tag{16}$$

We shall now assume that the <u>observed</u> deviations for the different observations (such as integration intervals) are uncorrelated. This means that

$$\langle \Delta y_j \Delta y_k \rangle = \sum_{\ell=1}^{N} \sum_{m=1}^{N} (T^{-1})_{j\ell} (T^{-1})_{km} \sum_{i=1}^{M} \langle \Delta A_i^2 \rangle \cdot \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_i}{\partial y_m} W_i^2 \tag{17}$$

We assume that the weights W_i are chosen in such a way that:

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/187 40
TELEX 8764 GEOFYSK S

$$\langle \Delta A_i^2 \rangle W_i = \text{constant} = \sigma^2 \tag{18}$$

We obtain:

$$\langle \Delta y_j \Delta y_k \rangle = \sigma^2 \sum_{\ell=1}^{N} \sum_{m=1}^{N} (T^{-1})_{j\ell} T_{\ell km} (T^{-1})_{km}$$

$$= \sigma^2 \sum_{\ell=1}^{N} T^{-1}{}_{j\ell} \cdot \delta_{\ell k} = \sigma^2 (T^{-1})_{jk} \tag{19}$$

Specifically the variances of the deviations become:

$$\langle \Delta y_j^2 \rangle = \sigma^2 (T^{-1})_{jj} \tag{20}$$

The absolute values of the weights are immaterial since they cancel in the final formula for th covariances.

Since we cannot know the variance of A_i in advance an estimate has to be made on the basis of the minimum value actually achieved for the sum (3) in the linear case.

The sum which was minimized is given by:

$$\varepsilon = \sum_{i=1}^{M} W_i (A_i - B_i (T^{-1}\vec{V}))^2 \tag{21}$$

where we have substituted the best estimate for $\vec{y}$ as obtained in (13). The quantity ε is a random variable and has an ensemble average different from zero because of the observational noise. Clearly the ensemble average becomes:

$$\langle \varepsilon \rangle = \sum_{i=1}^{M} W_i \langle (\Delta A_i - \sum_{\ell=1}^{N} \frac{\partial B_i}{\partial y_\ell} \cdot \Delta y_\ell)^2 \rangle \tag{22}$$

All the terms here can be expressed by the observational deviations ΔA_i:

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/187 40
TELEX 8764 GEOFYSK S

$$\langle \varepsilon \rangle = \sum_{i=1}^{M} W_i \left(\langle \Delta A_i^2 \rangle - 2 \sum_{\ell=1}^{N} \frac{\partial B_i}{\partial y_\ell} \langle \Delta A_i \Delta y_\ell \rangle + \sum_{\ell=1}^{N} \sum_{m=1}^{N} \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_i}{\partial y_m} \langle \Delta y_\ell \Delta y_m \rangle \right) \qquad (23)$$

The last term we determine from (19):

$$\sum_{i=1}^{M} W_i \sum_{\ell=1}^{N} \sum_{m=1}^{N} \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_i}{\partial y_m} (T^{-1})_{\ell m} \sigma^2$$

$$= \sigma^2 \sum_{\ell=1}^{N} \sum_{m=1}^{N} \left(\sum_{i=1}^{M} W_i \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_i}{\partial y_m} \right) (T^{-1})_{\ell m}$$

$$= \sigma^2 \sum_{\ell=1}^{N} \sum_{m=1}^{N} T_{\ell m} (T^{-1})_{m\ell} = \sigma^2 \sum_{\ell=1}^{N} 1 = N \cdot \sigma^2 \qquad (24)$$

The second term becomes, after substitution from (15):

$$2 \sum_{i=1}^{M} W_i \sum_{\ell=1}^{N} \frac{\partial B_i}{\partial y_\ell} \sum_{m=1}^{N} (T^{-1})_{\ell m} \sum_{j=1}^{M} \langle \Delta A_j \Delta A_i \rangle \frac{\partial B_j}{\partial y_m} W_j$$

$$= 2 \sum_{i=1}^{M} W_i \sum_{\ell=1}^{N} \frac{\partial B_i}{\partial y_\ell} \sum_{m=1}^{N} (T^{-1})_{\ell m} W_i \langle \Delta A_i^2 \rangle \frac{\partial B_i}{\partial y_m}$$

$$= 2\sigma^2 \sum_{\ell=1}^{N} \sum_{m=1}^{N} (T^{-1})_{\ell m} \sum_{i=1}^{M} W_i \frac{\partial B_i}{\partial y_\ell} \frac{\partial B_i}{\partial y_m} = 2N\sigma^2 \qquad (25)$$

Finally the first term becomes:

$$\sum_{i=1}^{M} W_i \langle \Delta A_i^2 \rangle = M \cdot \sigma^2$$

The mean value of the estimate of ε therefore becomes:

$$\langle \varepsilon \rangle = (M-N)\sigma^2$$

or: $$\sigma^2 = \frac{\langle \varepsilon \rangle}{M-N} \qquad (26)$$

EISCAT SCIENTIFIC ASSOCIATION
S-981 01 KIRUNA 1, SWEDEN
TELEPHONE 0980/ 187 40
TELEX 8754 GEOFYSK S