

PARAMETER ESTIMATION OF LINEAR SYSTEMS IN THE ABSENCE OF PROCESS NOISE: (ii) the Maximum Likelihood method

1. INTRODUCTION

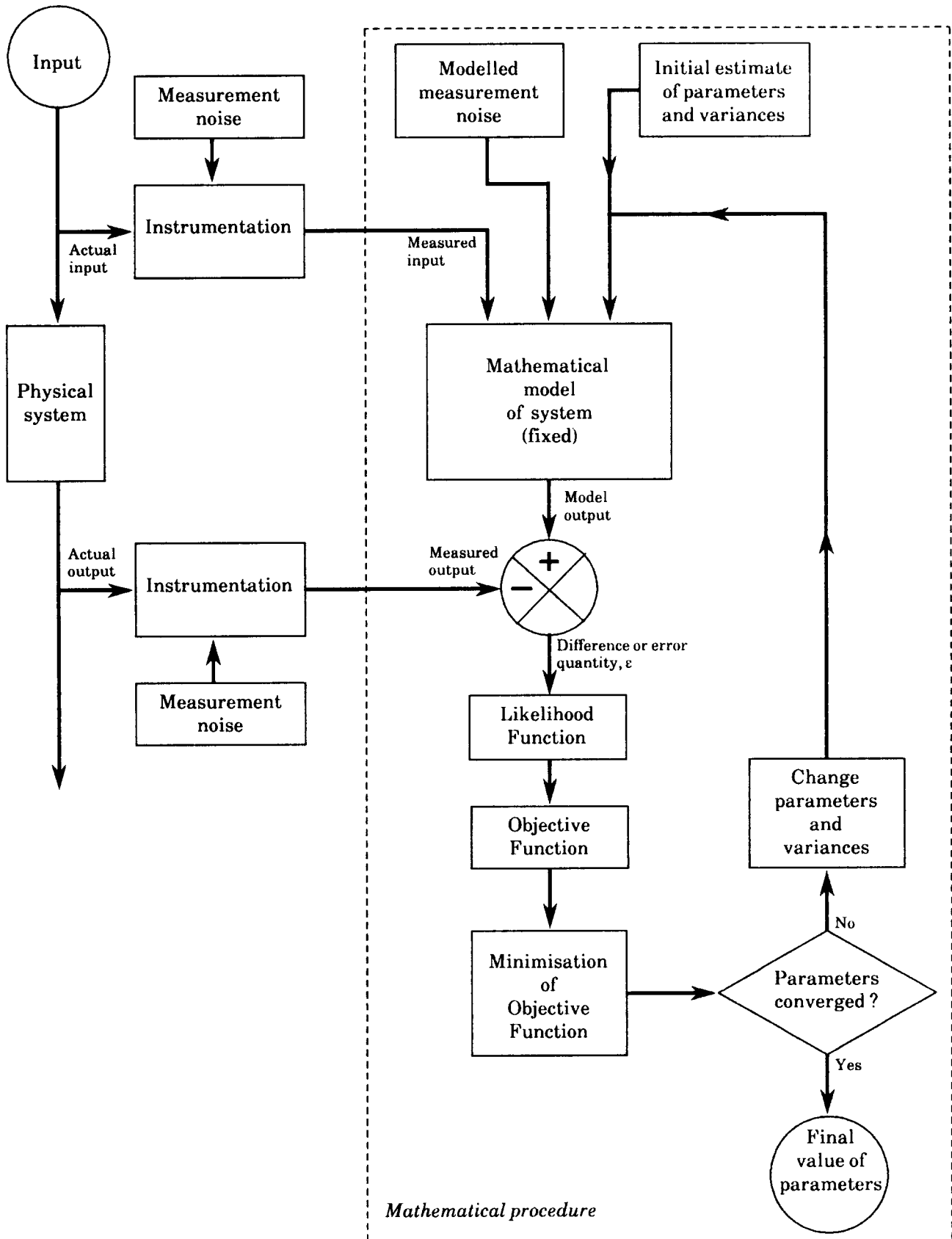
This is the second Data Item dealing with methods for the parameter estimation of linear systems.

An earlier Item, No. 87039 (Reference 1), introduced the parameter estimation process with methods based on the least-squares principle in which the system model was assumed free from process noise. The methods were essentially divided into direct, non-iterative methods in which the system equations were solved in terms of both input and output measurements and iterative methods that solved the system equations in terms of only input measurements. The former methods result in a decoupling of those equations containing different parameters so that each parameter is found separately. The latter methods result in a solution of all the parameters as a set and provide a measure of the correlation between pairs of parameters. A disadvantage of these latter methods is that biasing of the parameter estimates occurs when each of the measured state variables contains differing levels of noise.

This present Item considers the Maximum Likelihood method for linear systems whose models are assumed free from process noise. The method is iterative, like the latter class of method of Item No. 87039, but can also take account of differing levels of noise on each of the various instruments.

Sketch 1.1 shows a general flow chart of the Maximum Likelihood estimation process.

This Item briefly reviews the model of the system and the iterative parameter estimation methods of Item No. 87039 before introducing the Maximum Likelihood estimator, which is the keystone of the method. The estimation process is then described using a system model represented by two coupled first-order differential equations containing two unknown parameters. Two examples illustrate the application of this procedure. A further section describes the generalised method where the system model is represented by a multiparameter system of many degrees of freedom.



Sketch 1.1 General flow chart of Maximum Likelihood estimation process in the absence of process noise

2. NOTATION

Any coherent system of units may be used.

$\underline{A}$	plant or stability matrix
$\underline{B}$	control or input matrix
$\text{Cov}\{\dots\}$	covariance of $\{\dots\}$
$\underline{C}, \underline{D}, \underline{E}$	matrices associated with the observation (or measurement) equation
$\underline{F}, \underline{G}$	matrices associated with alternative form of observation equation
$\text{Det}\{\dots\}$	determinant of $\{\dots\}$
$E\{\dots\}$	expected value of $\{\dots\}$
h	total number of unknown parameters
I	identity matrix
J	Objective Function
$L\{\dots\}$	Likelihood Function of $\{\dots\}$
l	total number of output quantities
$\underline{M}$	information matrix
m	total number of coefficients associated with input quantities
N	total number of discrete time points
n	total number of coefficients associated with state variables
$p\{\dots\}$	probability density of $\{\dots\}$
$p\{y \gamma\}$	probability density of y given γ
q	modelled measurement noise
R	variance associated with difference quantity, ε
$\underline{S}$	covariance matrix
t	time
u	input variable
W	weighting variable
x	state variable

y	computed model quantity
z	measured quantity
γ	unknown parameter
Δ	denotes increment
δ_{ij}	Kronecker delta function, $= 1 \quad i = j$ $= 0 \quad i \neq j$
ε	difference between measured output quantity and model output quantity
θ	arbitrary random variable
σ	standard deviation
σ_e	estimate for standard deviation

Subscripts

I	denotes initial estimate
i	denotes discrete time points
max	maximum
min	minimum
0	initial condition
opt	optimised
s	denotes solution

Superscripts

T	denotes transpose
$'$	denotes an updated value

A bar below a quantity denotes a matrix or vector quantity.

One or two dots above a quantity denotes the first or second derivative, respectively, of that quantity with respect to time.

3. BACKGROUND

Before the Maximum Likelihood method for parameter estimation is described, a review of the system model and the least-squares methods will be given, together with a description of the Maximum Likelihood estimator.

3.1 Review of System Model

A previous Data Item, No. 87039 (Reference 1), presents the general form of a mathematical model of a linear dynamic system in which measurement noise only is present. The general state space representation of such a model, in which the elements of the coefficients are assumed time invariant, is given by

$$\dot{\underline{x}}(t) = \underline{A}\underline{x}(t) + \underline{B}\underline{u}(t), \quad \underline{x}(0) = \underline{x}_0 \quad (3.1a)$$

$$\text{and} \quad \underline{y}(t) = \underline{C}\underline{x}(t) + \underline{D}\dot{\underline{x}}(t) + \underline{E}\underline{u}(t) + \underline{q}(t), \quad (3.1b)$$

which are known as the state equations and observation (or measurement) equations, respectively. It is usual to call $\underline{A}$ the plant or stability matrix and $\underline{B}$ the input or control matrix. The observation equation, Equation (3.1b), is general, allowing independent observations of $\underline{x}(t)$ and $\dot{\underline{x}}(t)$. The elements of the coefficients $\underline{C}$, $\underline{D}$ and $\underline{E}$ are usually known. The minimum number of observation variables required is one for each degree of freedom, although additional outputs are desirable to give added accuracy and certainty.

The terms $\underline{q}(t)$ are included in the observation equation to allow for the possibility of modelling the measurement noise.

The vectors $\underline{x}(t)$, $\underline{u}(t)$, $\underline{y}(t)$ and $\underline{q}(t)$ in the above equations are the state vector of size $(n \times 1)$, the input vector of size $(m \times 1)$, the model output vector of size $(l \times 1)$ and the modelled measurement noise vector of size $(l \times 1)$, respectively, described by

$$\underline{x}(t) = [x(1, t), x(2, t), \dots, x(n, t)]^T, \quad (3.2a)$$

$$\underline{u}(t) = [u(1, t), u(2, t), \dots, u(m, t)]^T, \quad (3.2b)$$

$$\underline{y}(t) = [y(1, t), y(2, t), \dots, y(l, t)]^T \quad (3.2c)$$

$$\text{and} \quad \underline{q}(t) = [q(1, t), q(2, t), \dots, q(l, t)]^T. \quad (3.2d)$$

The unknown parameters to be found are some or all of the elements of $\underline{A}$ and/or $\underline{B}$ and are denoted by a row vector of size $(1 \times h)$, given by

$$\underline{\gamma} = [\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_h] \quad (3.3)$$

It may be noted that the number of output quantities, l , used in the analysis is entirely independent of the number of unknown parameters, h , to be found.

In an experimental test of the system, quantities, $\underline{z}_i$, corresponding to the observation quantities, $\underline{y}_i$, are measured at the discrete times, t_i , $i = 1, \dots, N$. The input quantities are also measured at these same discrete times, t_i , and are denoted by $\underline{z}(u)_i$.

The differences between the measured quantities and the observation quantities are expressed as

$$\underline{\varepsilon}_i = \underline{z}_i - \underline{y}_{si}, \quad (3.4)$$

where s denotes solution in terms of the input measurements, $\underline{z}(u)_i$. The observation equation solutions, $\underline{y}_{si}$, are obtained using the discrete form of Equation (3.1b), i.e.

$$\underline{y}_{si} = \underline{C}\underline{x}_{si} + \underline{D}\dot{\underline{x}}_{si} + \underline{E}\underline{z}(u)_i + \underline{q}_i, \quad (3.5)$$

where $\underline{x}_{si}$ are the solutions of the state equations in terms of the input measurements. It is permissible at this stage of the analysis to replace the term $\underline{D}\dot{\underline{x}}_{si}$, in Equation (3.5), by a discrete form solution of the state Equation (3.1a). In this case Equation (3.5) may be rewritten as

$$\underline{y}_{si} = \underline{F}\underline{x}_{si} + \underline{G}\underline{z}(u)_i + \underline{q}_i, \quad (3.6)$$

which is now seen to contain only the state solutions and the input measurements.

3.2 Review of Least-Squares Parameter Estimation

Item No. 87039 introduces the parameter estimation process with methods based on the least-squares principle, in which attention is restricted to linear system models that are assumed free from process noise. The methods are classified according to whether the system model equations are solved in terms of both input and output measurements or whether the equations are solved in terms of the input measurements only. The former methods lead to a direct solution for the unknown parameters while the latter methods involve an iterative procedure starting from chosen estimates for each of the parameters. Within these two classes, a range of methods is available depending on the measured outputs of the system, i.e. whether the state variables, $\underline{x}_i$, are measured, whether the derivatives of the state variables with respect to time, $\dot{\underline{x}}_i$, are measured or whether $\underline{x}_i$ and $\dot{\underline{x}}_i$ are measured as a combination.

One of the iterative methods included in Item No. 87039 that particularly features in the parameter estimation literature is the Output Error method, also referred to as the generalised least-squares or weighted least-squares method. This method together with its variants, also included in Item No. 87039, has a close affinity with the Maximum Likelihood method.

From measurements of either $\underline{x}_i$ and $\underline{u}_i$ (the inputs) or $\underline{x}_i$ and $\dot{\underline{x}}_i$ as a combination and $\underline{u}_i$, an Objective Function, J , is formed based on the square of the differences, $\underline{\varepsilon}_i$, between the measured and computed (model) values. The Objective Function takes the form

$$J = W(1) \sum_{i=1}^N \varepsilon(1)_i^2 + W(2) \sum_{i=1}^N \varepsilon(2)_i^2 + \dots + W(l) \sum_{i=1}^N \varepsilon(l)_i^2, \quad (3.7)$$

where l is the total number of output quantities and $W(1), W(2), \dots, W(l)$ are appropriately chosen constant weightings primarily to make each term in Equation (3.7) of comparable magnitude.

This method results in an iterative solution for the unknown parameters involving the complete set of state equations of the model of the system. The methods result in biased estimates when measurement noise exists in the input variables. Since the weightings associated with each of the terms of the Objective Function are fixed throughout the analysis, further biasing of the parameter estimates results when each of the measured state variables contains differing levels of noise. This is often the case since widely different

types of transducers have to be used to measure the different types of output. The choice of the values of $W(1)$, $W(2)$, ..., $W(I)$ is then difficult.

3.3 The Maximum Likelihood Estimator

The Maximum Likelihood method can take account of the differing levels of noise on the various instruments used for obtaining the measurements and therefore accommodates to some extent the disadvantages inherent in the iterative least-squares methods. In this method the equivalent weightings are determined, rather than being specified.

A Maximum Likelihood estimator can be introduced in the following way. Consider a set of N samples, $\theta_1, \theta_2, \dots, \theta_N$, of a single random variable, $\{\theta|\underline{\gamma}\}$, which depends on a number of parameters, $\underline{\gamma}$, where

$$\underline{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_h]. \quad (3.8)$$

At optimum values of the parameter, $\underline{\gamma}_{opt}$, the joint probability density function is assumed to be known and denoted by

$$p_N\{\theta\} = p_N\{\theta_1, \theta_2, \dots, \theta_N | \underline{\gamma}_{opt}\}. \quad (3.9)$$

When the samples are uncorrelated the density function may be rewritten as

$$p_N\{\theta\} = p\{\theta_1 | \underline{\gamma}_{opt}\} p\{\theta_2 | \underline{\gamma}_{opt}\} \dots p\{\theta_N | \underline{\gamma}_{opt}\}, \quad (3.10)$$

or, more concisely, as

$$p_N\{\theta\} = \prod_{i=1}^N p\{\theta_i | \underline{\gamma}_{opt}\}. \quad (3.11)$$

A Likelihood Function, $L\{\underline{\gamma}\}$, is formulated as

$$L\{\underline{\gamma}\} = \prod_{i=1}^N p\{\theta_i | \underline{\gamma}\}, \quad (3.12)$$

where $p\{\theta_i | \underline{\gamma}\}$ is the same function as $p\{\theta_i | \underline{\gamma}_{opt}\}$ but where the set of samples, θ_i , $i = 1, \dots, N$, is based on non-optimum values of the parameter, $\underline{\gamma}$.

When this Likelihood Function is maximised with respect to the parameters then

$$\underline{\gamma} \rightarrow \underline{\gamma}_{opt}, \quad (3.13)$$

$$p\{\theta_i | \underline{\gamma}\} \rightarrow p\{\theta_i | \underline{\gamma}_{opt}\} \quad (3.14)$$

$$\text{and hence } L\{\underline{\gamma}\} \rightarrow L_{max}\{\underline{\gamma}_{opt}\} = \prod_{i=1}^N p\{\theta_i | \underline{\gamma}_{opt}\}. \quad (3.15)$$

This last equation is called the Maximum Likelihood Equation and the solutions, $\underline{\gamma}_{opt}$, that satisfy this equation are called the Maximum Likelihood estimates[†] for $\underline{\gamma}$.

There may be an element of arbitrariness in the definition of the Likelihood Function of Equation (3.12); however, the process of maximising the Likelihood Function provides the unique set of optimised values of the parameters.

4. TWO UNKNOWN PARAMETERS IN A SECOND-ORDER SYSTEM

The Maximum Likelihood method is based on finding the set of optimised parameters of a system which maximises a Likelihood Function. At this optimised condition the Maximum Likelihood Function is the joint probability density function of the observations.

The Maximum Likelihood method is illustrated by considering a system composed of two coupled first-order differential equations containing two unknown parameters, given by

$$\dot{x}(1, t) = x(2, t), \quad x(1, 0) = 0 \quad (4.1a)$$

$$\text{and} \quad \dot{x}(2, t) = \gamma_1 x(1, t) + \gamma_2 x(2, t) + u(t), \quad x(2, 0) = 0, \quad (4.1b)$$

where $x(1, t)$ and $x(2, t)$ are the state variables of the system, $u(t)$ is the input to the system and γ_1 and γ_2 are the unknown parameters to be found.

4.1 Direct Measurements of State Variables are Available

It is assumed in this sub-section, for simplicity, that both state variables are directly measured and so the observation equations in discrete form are

$$y(1)_i = x(1)_i, \quad i = 1, 2, \dots, N \quad (4.2a)$$

$$\text{and} \quad y(2)_i = x(2)_i, \quad i = 1, 2, \dots, N, \quad (4.2b)$$

where i denotes the discrete values of t_i at which the measurements are taken. The q_i terms are absent from Equations (4.2a) and (4.2b) since the measurement noise is not specifically modelled.

Measurements of the observation variables and input variables are denoted by

$$z(1)_i = \text{the measured values of } y(1)_i, \quad (4.3a)$$

$$z(2)_i = \text{the measured values of } y(2)_i \quad (4.3b)$$

$$\text{and} \quad z(u)_i = \text{the measured values of } u_i. \quad (4.3c)$$

The differences between the measured observation variables, $z(1)_i$ and $z(2)_i$, and the computed values of the observation variables, $y_s(1)_i$ and $y_s(2)_i$, in terms of the measured input, are denoted by

$$\varepsilon(1)_i = z(1)_i - y_s(1)_i \equiv z(1)_i - x_s(1)_i \quad (4.4a)$$

$$\text{and} \quad \varepsilon(2)_i = z(2)_i - y_s(2)_i \equiv z(2)_i - x_s(2)_i, \quad (4.4b)$$

[†] The symbol $\hat{\underline{\gamma}}$ is often used to denote the Maximum Likelihood estimates for $\underline{\gamma}$.

where the suffix s denotes solution of the state equations in terms of the input measurements, $z(u)_i$, for given values of γ_1 and γ_2 .

In the Maximum Likelihood method it is assumed that the measurement noise associated with each of the instruments measuring $y(1)$ and $y(2)$ is white with a Gaussian probability distribution and a mean level of zero. It is also assumed in the method that when an optimised solution has been obtained, *i.e.* when the unknown parameters have their optimised values, then the differences, $\epsilon_{opt}(1)_i$ and $\epsilon_{opt}(2)_i$, are samples of the measurement noise. Consequently the mean and variances of these differences may be expressed as

$$E\{\epsilon_{opt}(1)_i\} = \frac{1}{N} \sum_{i=1}^N \epsilon_{opt}(1)_i \rightarrow 0, \quad (4.5a)$$

$$E\{\epsilon_{opt}(2)_i\} = \frac{1}{N} \sum_{i=1}^N \epsilon_{opt}(2)_i \rightarrow 0, \quad (4.5b)$$

$$\frac{1}{N} \sum_{i=1}^N [\epsilon_{opt}(1)_i]^2 \rightarrow [\sigma_{opt}(1)]^2 \quad (4.6a)$$

$$\text{and} \quad \frac{1}{N} \sum_{i=1}^N [\epsilon_{opt}(2)_i]^2 \rightarrow [\sigma_{opt}(2)]^2, \quad (4.6b)$$

where $\sigma_{opt}(1)$ and $\sigma_{opt}(2)$ are approximations to the standard deviations of the measurement noise of the instruments used to measure $z(1)_i$ and $z(2)_i$, respectively.

At the optimised conditions for each error, $\epsilon_{opt}(1)_i$, the probability density function is given by

$$p\{\epsilon_{opt}(1)_i | \gamma_{1opt}, \gamma_{2opt}\} = \frac{1}{[2\pi\sigma_{opt}(1)^2]^{1/2}} \exp\left[-\frac{\epsilon_{opt}(1)_i^2}{2\sigma_{opt}(1)^2}\right]. \quad (4.7)$$

The joint probability density function of the sequence of errors $\epsilon_{opt}(1)_i$, $i = 1$ to N , and $\epsilon_{opt}(2)_i$, $i = 1$ to N , assuming all the errors are uncorrelated, is given (see Section 3.3) by

$$p\{|\} = \left| \prod_{i=1}^N \frac{1}{[2\pi\sigma_{opt}(1)^2]^{1/2}} \exp\left[-\frac{\epsilon_{opt}(1)_i^2}{2\sigma_{opt}(1)^2}\right] \right| \left| \prod_{i=1}^N \frac{1}{[2\pi\sigma_{opt}(2)^2]^{1/2}} \exp\left[-\frac{\epsilon_{opt}(2)_i^2}{2\sigma_{opt}(2)^2}\right] \right|, \quad (4.8)$$

or, simplifying,

$$p\{|\} = \frac{1}{[2\pi]^N [\sigma_{opt}(1)^2 \sigma_{opt}(2)^2]^{N/2}} \exp\left[-\frac{1}{2\sigma_{opt}(1)^2} \sum_{i=1}^N \epsilon_{opt}(1)_i^2 - \frac{1}{2\sigma_{opt}(2)^2} \sum_{i=1}^N \epsilon_{opt}(2)_i^2\right]. \quad (4.9)$$

The Maximum Likelihood Function is

$$L_{max}\{\gamma_{opt}\} = p\{\cdot\}, \quad (4.10)$$

where $p\{\cdot\}$ is given by Equation (4.9).

A Likelihood Function may now be formulated, based on the form of Equations (4.9) and (4.10), as

$$L\{\gamma\} = \frac{1}{[2\pi]^N [\sigma(1)^2 \sigma(2)^2]^{N/2}} \exp \left[-\frac{1}{2\sigma(1)^2} \sum_{i=1}^N \varepsilon(1)_i^2 - \frac{1}{2\sigma(2)^2} \sum_{i=1}^N \varepsilon(2)_i^2 \right], \quad (4.11)$$

where $\sigma(1)$, $\sigma(2)$, $\varepsilon(1)$ and $\varepsilon(2)$ are new variables that are based on non-optimum values of the parameters γ_1 and γ_2 .

As the optimisation proceeds $L\{\gamma\}$ is maximised and each variable approaches its optimised value, *i.e.*

$$\sigma(1)^2 \rightarrow \sigma_{opt}(1)^2,$$

$$\sigma(2)^2 \rightarrow \sigma_{opt}(2)^2,$$

$$\gamma_1 \rightarrow \gamma_{1opt}$$

and $\gamma_2 \rightarrow \gamma_{2opt}.$

In the Maximum Likelihood method it is usual to define the Objective Function, J (often referred to as the Cost Function), as the negative logarithm of the Likelihood Function and so

$$J = -\ln \{L\{\gamma\}\}, \quad (4.12a)$$

or

$$J = \frac{1}{2\sigma(1)^2} \sum_{i=1}^N \varepsilon(1)_i^2 + \frac{1}{2\sigma(2)^2} \sum_{i=1}^N \varepsilon(2)_i^2 + \frac{N}{2} \ln \{\sigma(1)^2\} + \frac{N}{2} \ln \{\sigma(2)^2\} + \text{constant}. \quad (4.12b)$$

This equation may be rewritten using Equations (4.4a) and (4.4b) as

$$J = \frac{1}{2\sigma(1)^2} \sum_{i=1}^N [z(1)_i - x_s(1)_i]^2 + \frac{1}{2\sigma(2)^2} \sum_{i=1}^N [z(2)_i - x_s(2)_i]^2 + \frac{N}{2} \ln \{\sigma(1)^2\} + \frac{N}{2} \ln \{\sigma(2)^2\} + \text{constant}. \quad (4.12c)$$

The maximum value of the Likelihood Function, L , occurs when the Objective Function, J , is a minimum. It may be noted that the 'coefficients', $1/(2\sigma(1)^2)$ and $1/(2\sigma(2)^2)$, in Equation (4.12c) are similar to the weightings, $W(1)$ and $W(2)$, in Equation (3.7).

The Objective Function, J , can be regarded as a function of the four independent unknowns, $\sigma(1)^2$, $\sigma(2)^2$,

γ_1 and γ_2 . The minimum value of J is at the optimised condition and so

$$\left\{ \frac{\partial J}{\partial \sigma(1)^2} \right\}_{opt} = \left\{ \frac{\partial J}{\partial \sigma(2)^2} \right\}_{opt} = \left\{ \frac{\partial J}{\partial \gamma_1} \right\}_{opt} = \left\{ \frac{\partial J}{\partial \gamma_2} \right\}_{opt} = 0. \quad (4.13)$$

The first two equations give

$$\{\sigma(1)^2\}_{opt} = \frac{1}{N} \sum_{i=1}^N \varepsilon_{opt}(1)_i^2 \quad (4.14a)$$

$$\text{and} \quad \{\sigma(2)^2\}_{opt} = \frac{1}{N} \sum_{i=1}^N \varepsilon_{opt}(2)_i^2, \quad (4.14b)$$

which satisfy Equations (4.6a) and (4.6b).

The second two equations give

$$\left| \frac{1}{\sigma_{opt}(1)^2} \sum_{i=1}^N [z(1)_i - x_s(1)_i] \frac{\partial x_s(1)_i}{\partial \gamma_1} + \frac{1}{\sigma_{opt}(2)^2} \sum_{i=1}^N [z(2)_i - x_s(2)_i] \frac{\partial x_s(2)_i}{\partial \gamma_1} \right|_{\substack{\gamma_1 = \gamma_{1opt} \\ \gamma_2 = \gamma_{2opt}}} = 0 \quad (4.15a)$$

$$\text{and} \quad \left| \frac{1}{\sigma_{opt}(1)^2} \sum_{i=1}^N [z(1)_i - x_s(1)_i] \frac{\partial x_s(1)_i}{\partial \gamma_2} + \frac{1}{\sigma_{opt}(2)^2} \sum_{i=1}^N [z(2)_i - x_s(2)_i] \frac{\partial x_s(2)_i}{\partial \gamma_2} \right|_{\substack{\gamma_1 = \gamma_{1opt} \\ \gamma_2 = \gamma_{2opt}}} = 0. \quad (4.15b)$$

To solve these latter two nonlinear equations it is necessary to adopt an iterative approach. The procedure that will be used is a second-order optimisation procedure known as the Modified Newton-Raphson procedure or Newton-Gauss procedure. This procedure approximates the Objective Function by the first three terms of its Taylor series expansion about initial estimates of the parameters, say γ_{I1} and γ_{I2} . Expanding J in turn about γ_{I1} and γ_{I2} gives the following two equations

$$J(\gamma_1, \gamma_{I2}) = J(\gamma_{I1}, \gamma_{I2}) + (\gamma_1 - \gamma_{I1}) \frac{\partial J}{\partial \gamma_1} + \frac{(\gamma_1 - \gamma_{I1})^2}{2} \frac{\partial^2 J}{\partial \gamma_1^2} \quad (4.16a)$$

$$\text{and} \quad J(\gamma_{I1}, \gamma_2) = J(\gamma_{I1}, \gamma_{I2}) + (\gamma_2 - \gamma_{I2}) \frac{\partial J}{\partial \gamma_2} + \frac{(\gamma_2 - \gamma_{I2})^2}{2} \frac{\partial^2 J}{\partial \gamma_2^2}, \quad (4.16b)$$

where γ_1 and γ_2 are related to increments in the parameters, $\Delta\gamma_1$ and $\Delta\gamma_2$, by

$$\gamma_1 = \gamma_{I1} + \Delta\gamma_1 \quad (4.17a)$$

$$\text{and} \quad \gamma_2 = \gamma_{I2} + \Delta\gamma_2. \quad (4.17b)$$

Equations (4.16a) and (4.16b) are now differentiated with respect to the unknown parameters γ_1 and γ_2

to give, upon rearrangement and with the use of Equations (4.17a) and (4.17b),

$$\Delta\gamma_1 = -\left\{\frac{\partial^2 J}{\partial \gamma_1^2}\right\}^{-1} \left\{\frac{\partial J}{\partial \gamma_1}\right\} \quad (4.18a)$$

and

$$\Delta\gamma_2 = -\left\{\frac{\partial^2 J}{\partial \gamma_2^2}\right\}^{-1} \left\{\frac{\partial J}{\partial \gamma_2}\right\}. \quad (4.18b)$$

The first and second derivatives of $J(\gamma_{I1}, \gamma_{I2})$ may be found from Equation (4.12c) as

$$\frac{\partial J}{\partial \gamma_j} = -\frac{1}{\sigma(1)^2} \sum_{i=1}^N [z(1)_i - x_s(1)_i] \frac{\partial x_s(1)_i}{\partial \gamma_j} - \frac{1}{\sigma(2)^2} \sum_{i=1}^N [z(2)_i - x_s(2)_i] \frac{\partial x_s(2)_i}{\partial \gamma_j} \quad (4.19a)$$

and

$$\begin{aligned} \frac{\partial^2 J}{\partial \gamma_j^2} = & \frac{1}{\sigma(1)^2} \sum_{i=1}^N \left[\frac{\partial x_s(1)_i}{\partial \gamma_j} \right]^2 + \frac{1}{\sigma(2)^2} \sum_{i=1}^N \left[\frac{\partial x_s(2)_i}{\partial \gamma_j} \right]^2 \\ & + \frac{1}{\sigma(1)^2} \sum_{i=1}^N [z(1)_i - x_s(1)_i] \frac{\partial^2 x_s(1)_i}{\partial \gamma_j^2} + \frac{1}{\sigma(2)^2} \sum_{i=1}^N [z(2)_i - x_s(2)_i] \frac{\partial^2 x_s(2)_i}{\partial \gamma_j^2}, \end{aligned} \quad (4.19b)$$

where $j = 1, 2$ covers the two cases for the parameters γ_1 and γ_2 . The above equations represent the second-order Newton-Raphson procedure. A modified form, known as the Newton-Gauss procedure is obtained by dropping the second two smaller terms of Equation (4.19b) (see Reference 3).

A solution procedure can now be established. With the initial values of γ_{I1} , γ_{I2} , $\sigma(1)_I$ and $\sigma(2)_I$, the increments in the parameters $\Delta\gamma_1$ and $\Delta\gamma_2$ may be found from Equations (4.18a), (4.18b), (4.19a) and (4.19b). From these increments an updated set of parameters, denoted by a dash, is obtained from

$$\gamma_1' = \gamma_{I1} + \Delta\gamma_1 \quad (4.20a)$$

and

$$\gamma_2' = \gamma_{I2} + \Delta\gamma_2. \quad (4.20b)$$

These parameters together with Equations (4.18a), (4.18b), (4.19a) and (4.19b) are used to obtain a further set of parameters and this iteration process is repeated until a minimum of the Objective Function is obtained, *i.e.* when $\Delta\gamma_1$ and $\Delta\gamma_2$ are sufficiently small and γ_1 and γ_2 approach their optimised values.

At each stage of the iteration process, once the increments in the parameters have been obtained, the 'variances' are also updated from

$$\{\sigma(1)'\}^2 = \frac{1}{N} \sum_{i=1}^N \{\epsilon(1)_i'\}^2 \quad (4.21a)$$

and

$$\{\sigma(2)'\}^2 = \frac{1}{N} \sum_{i=1}^N \{\epsilon(2)_i'\}^2, \quad (4.21b)$$

and consequently at the optimised solution the variances also attain their optimised values. The first and