

Parameter estimation of linear systems in the absence of process noise: (i) methods based on the least-squares principle

Endorsed by
The Royal Aeronautical Society

ESDU DATA ITEMS

Data Items provide validated information in engineering design and analysis for use by, or under the supervision of, professionally qualified engineers. The data are founded on an evaluation of all the relevant information, both published and unpublished, and are invariably supported by original work of ESDU staff engineers or consultants. The whole process is subject to independent review for which crucial support is provided by industrial companies, government research laboratories, universities and others from around the world through the participation of some of their leading experts on ESDU Technical Committees. This process ensures that the results of much valuable work (theoretical, experimental and operational), which may not be widely available or in a readily usable form, can be communicated concisely and accurately to the engineering community.

We are constantly striving to develop new work and review data already issued. Any comments arising out of your use of our data, or any suggestions for new topics or information that might lead to improvements, will help us to provide a better service.

THE PREPARATION OF THIS DATA ITEM

The work on this particular Data Item was monitored and guided by the Dynamics Committee which has the following constitution:

Chairman

Dr H.H.B.M. Thomas — Independent

Members

Prof. G.T.S. Done	— The City University, London
Prof. G.J. Hancock	— Queen Mary College, London University
Mr M.R. Heath	— British Aerospace plc, Army Weapons Division, Stevenage
Dr E.M. James	— The University College of Wales, Aberystwyth
Mr M.R. Rushton	— British Aerospace plc, Military Aircraft Division, Kingston
Dr M. Tobak*	— NASA, Ames Research Centre, California, USA
Dr M.A. Woodhead	— University of Salford,

* Corresponding Member.

with further assistance from Dr A.J. Ross of the Royal Aircraft Establishment, Farnborough.

The technical work involved in the assessment of the available information and the construction and subsequent development of the Data Item was undertaken by

Mr C.J. Loughton — Head of Mechanical Motion and System Dynamics Group.

The Dynamics Committee, which first met in January 1962, has the co-operation of many engineers and scientists in industry, research establishments and universities from whom much assistance and unpublished information is being received.

PARAMETER ESTIMATION OF LINEAR SYSTEMS IN THE ABSENCE OF PROCESS NOISE: (i) methods based on the least-squares principle

CONTENTS

	Page
1. INTRODUCTION	1
2. NOTATION	3
3. SYSTEM MODEL	5
4. SINGLE UNKNOWN PARAMETER IN A SINGLE FIRST-ORDER SYSTEM	7
4.1 Measurement of State Variable and Input Variable	7
4.1.1 Direct solution for unknown parameter	8
4.1.2 Iterative solution for unknown parameter (Output Error method)	9
4.1.3 Form of the input	12
4.1.4 Example	12
4.2 Measurement of State Variable Rate and Input Variable	17
4.2.1 Direct solution for unknown parameter	17
4.2.2 Iterative solution for unknown parameter	18
4.2.3 Example	19
4.3 Measurement of State Variable Rate, State Variable and Input Variable (Equation Error Method)	23
4.3.1 Example	24
4.4 Instrumentation Calibration	25
5. GENERALISED METHODS	26
5.1 Measurement of State Variables (or Equivalent Output Variables) and Input Variables	27
5.1.1 Direct solution for unknown parameters	27
5.1.2 Iterative solution for unknown parameters (Output Error method)	29
5.2 Measurement of Rate of Change of State Variables (or Equivalent Output Variables) and Input Variables	33
5.2.1 Direct solution for unknown parameters	34
5.2.2 Iterative solution for unknown parameters	36
5.3 Measurements of the State Variable Rate, State Variables and Input Variables (Equation Error Method)	37
6. REFERENCES	39

PARAMETER ESTIMATION OF LINEAR SYSTEMS IN THE ABSENCE OF PROCESS NOISE: (i) methods based on the least-squares principle

1. INTRODUCTION

Identification, in the context of the response of a system, is the study of the correlation between the actual behaviour or performance of a real system and the predicted behaviour from a mathematical model which is assumed to characterise that system. When the mathematical model is expressed in terms of a known set of differential equations with constant coefficients[†], in which a number of these coefficients are unknown, the parameter estimation process aims at determining the best values of these unknown coefficients from the measured input and output of the system. Output measurements differ from the predicted model outputs for various reasons: firstly, as a result of an imperfect model representation of the real system and, secondly, as the result of noise, of which there are two kinds, process noise and measurement noise. Process noise affects the actual output of a system and can be an input (or external noise), such as the turbulence affecting the motion of an aircraft in flight, or an internal system quantity, such as the noise introduced by a sensor that forms part of a feedback loop of a system. Measurement noise arises from noise in the instrumentation used to measure the input and output. Process noise affects the actual output, whereas measurement noise does not. The remaining discrepancy between measured quantities and those deduced from the mathematical model may be termed observation error.

This Item restricts attention to systems that are assumed free from process noise and whose models are assumed linear.

This Data Item presents a number of digital optimisation techniques for parameter estimation that are based on the least-squares principle, a technique known mainly from its use in curve fitting or regression analysis. These digital computer based methods have superseded, to a large extent, older techniques such as steady-state tests, the time vector method and response matching using an analogue computer, since these earlier methods were dependent on the operator's skill and judgement in the application of these methods and on the restrictive nature of the form and magnitude of the inputs that could be accommodated.

The essence of techniques described here is the minimisation of a chosen Objective Function (often referred to as a Cost Function) that is formed from the square of the difference between the measured output of a particular variable and the model output of this same variable, where the model output is calculated in terms of the unknown coefficients for the same input. Broadly the techniques divide into the following two classes:

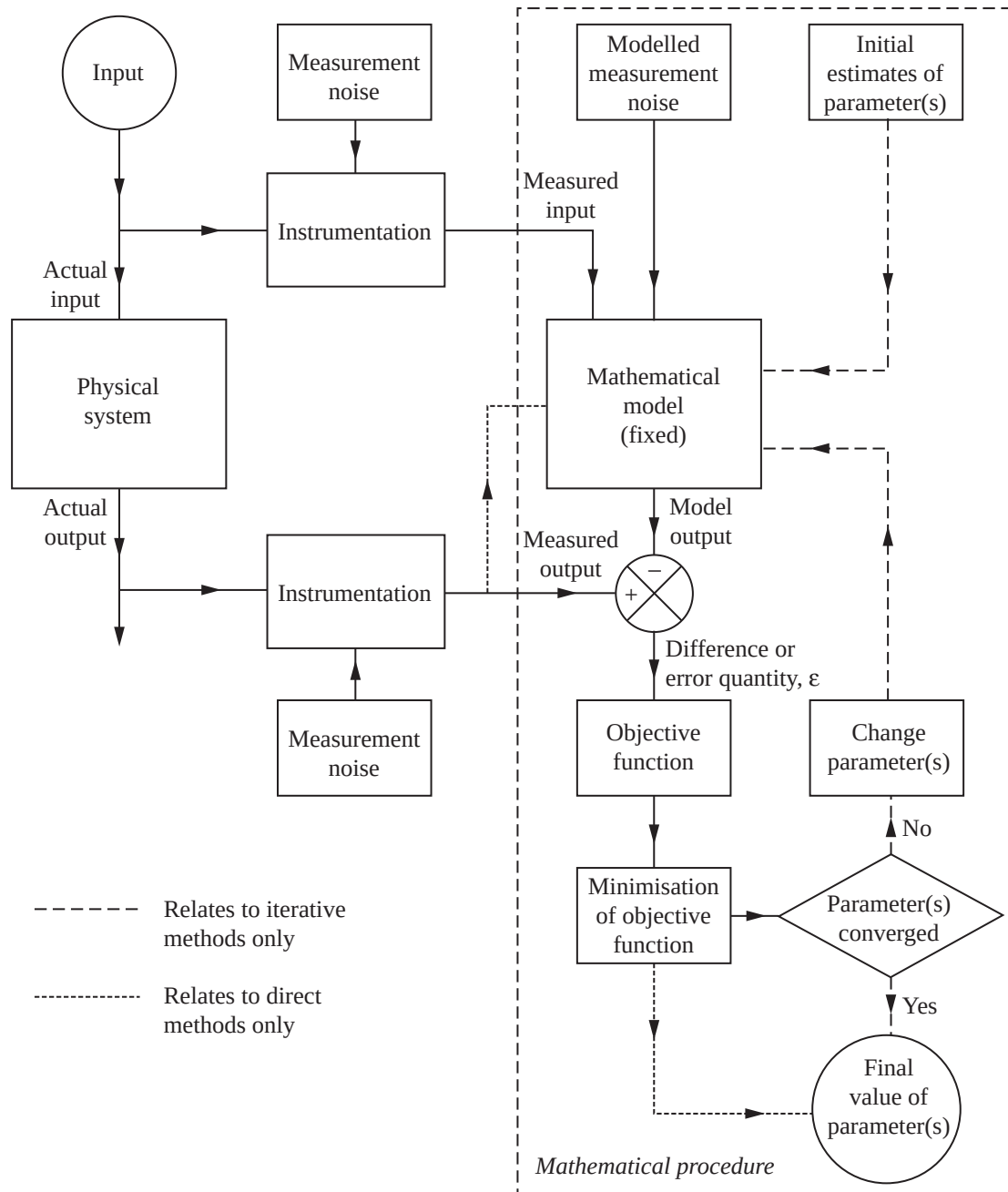
- (i) the system equation(s) are solved in terms of all the measurements, both input and output,
- (ii) the system equation(s) are solved in terms of the input measurements only.

Within each of the above two classes the methods vary according to the availability of the measured outputs of the system. Sketch 1.1, where it is assumed that only measurement noise is present in the system, shows a general flow chart of the parameter estimation process.

Section 3 of this Item presents the general form of the mathematical model of a linear system in which only measurement noise is assumed present within the model of the system. Section 4 introduces various parameter estimation techniques that consider a system model described by a single first-order equation involving a single unknown parameter. Two of these techniques are usually referred to in the literature as the Equation Error method and the Output Error method. An example illustrates the application of each of

[†] In the literature "parameter" is a commonly used term for those coefficients in the analysis that are assumed unknown.

these techniques. Section 5 considers the generalised form of the equations representing a multiparameter system of many degrees-of-freedom. References 1 to 9 provide sources of information supplementary to that in this Item.



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

Another class of method that may be used for parameter estimation is the Maximum Likelihood method. It is an advanced method capable of dealing with both linear and nonlinear systems containing measurement noise and is covered in other Items in this series (See References 10 to 14). For systems with process noise, the Filter Error method is used (see References 3, 11 and 12).

2. NOTATION

$\underline{A}$	plant or stability matrix
$\underline{\hat{A}}, \underline{\hat{A}}1$	row vectors (used for coefficients when state and input variables are combined)
a	coefficient of state variable (used for system approximated by single first-order differential equation with single unknown parameter)
$\underline{B}$	control or input matrix
$\text{Cov}\{\dots\}$	covariance of $\{\dots\}$
$\underline{C}, \underline{D}$	matrices associated with the observation (or measurement) equation
$E\{\dots\}$	expected value of $\{\dots\}$
$\underline{E}, \underline{F}, \underline{G}$	matrices associated with alternative form of observation equation
$\underline{F}$	residual error matrix
$\underline{G}$	sensitivity matrix containing partial derivatives of calculated outputs with respect to unknown parameters
$\underline{H}$	parameter increment matrix
h	total number of unknown parameters
I	identity matrix
J	Objective Function
k	calibration constant
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
q	measurement noise
$\underline{S}$	covariance matrix
t	time
u	input variable

W	weighting parameter
$\underline{X}$	vector combining state variable and input variable (used in generalised methods)
x	state variable
$\underline{Y}$	vector containing all measured outputs
y	computed model quantity
z	measured quantity
γ	unknown parameter
$\hat{\gamma}$	least-squares estimator for γ
Δ	denotes increment
ε	difference between measured output quantity and model output quantity
σ	standard deviation
σ_e	estimate for standard deviation

Subscripts

I	denotes initial estimate
i	denotes discrete time points
x, u	pertaining to state quantity and input quantity, respectively

Superscripts

T	denotes transpose
$'$	denotes updated value

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

A dot above a variable denotes differentiation of that variable with respect to time.

3. SYSTEM MODEL

This section presents the general form of the mathematical model of a linear system in which only measurement noise is assumed present. The model consists on the one part as a set of differential equations with constant coefficients the solution of which is known to represent reasonably the behaviour of the real system, and on the other as a set of modelled outputs relating the theoretical counterparts of the measured output variables with the state variables within the differential equations.

The general state space representation of a model of a linear system, in which the elements of the coefficients are assumed time invariant, is given by the following equations

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

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

which are known as the state equations and observation (or measurement) equations, respectively. The differential equations, Equations (3.1a), represent the equations governing the physical system in question and Equations (3.1b) define the relationships between the model outputs (the theoretical counterparts of the measured variables) and the state variables. In Equations (3.1a) it is usual to call $\underline{A}$ the plant or stability matrix and $\underline{B}$ the input or control matrix.

The vectors $\underline{x}(t)$, $\underline{u}(t)$, $\underline{y}(t)$ and $\underline{q}(t)$ in the above equations are, respectively, 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 $(l \times 1)$, described by

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

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

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

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

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

The unknowns of the system, commonly referred to as “parameters”, which are to be found, are some, or all, of the coefficients of Equations of (3.1a), i.e. some, or all, of the elements of the matrices $\underline{A}$ and $\underline{B}$.

It may be noted that when the observation equation is expressed in the form shown in Equations (3.1b), some, or all, of the elements of $\underline{C}$ and $\underline{D}$ may involve the elements of $\underline{A}$ and $\underline{B}$, respectively, where it should be recognised that no new unknown parameter is present in $\underline{C}$ or $\underline{D}$.

Equations (3.1b) may, in some cases, occur in the following form

$$\underline{y}(t) = \underline{E} \underline{x}(t) + \underline{F} \underline{u}(t) + \underline{G} \dot{\underline{x}}(t) + \underline{q}(t), \quad (3.2)$$

which contains the additional term involving the derivatives of the state variables with respect to time. In this case the matrices $\underline{E}$, $\underline{F}$ and $\underline{G}$ would be independent of $\underline{A}$ and $\underline{B}$ and would be completely known. Equations (3.2) can always be expressed in the form adopted in Equations (3.1b) by using a substitution of the state equation.

In an experimental test of the system, measured quantities, $\underline{z}_i$, corresponding to the calculated model quantities, $\underline{y}_i$, at the discrete times, t_i , $i = 1, N$ are available for a measured change in the input vector, $\underline{u}_i$. Equations (3.1a) and (3.1b) are thus replaced by

$$\dot{\underline{x}}(t) = \underline{A} \underline{x}(t) + \underline{B} \underline{u}(t) \quad (3.3a)$$

$$\text{and} \quad \underline{y}_i = \underline{C} \underline{x}_i + \underline{D} \underline{u}_i + \underline{q}_i, \quad i = 1, 2, \dots, N, \quad (3.3b)$$

where i is an integer denoting the individual values of t_i .

The values of the parameters, *i.e.* the unknown elements of $\underline{A}$ and $\underline{B}$, are obtained by minimising an Objective Function, J , with respect to each of the parameters in turn, formed from the square of the difference between the measured quantities, $\underline{z}_i$, and the predicted quantities, $\underline{y}_i$. Typically the Objective Function is formed from terms of the form

$$J = \sum_{i=1}^N \varepsilon_i^2 = \sum_{i=1}^N (z_i - y_i)^2, \quad (3.4)$$

where ε_i is the difference quantity. Denoting the unknown parameters by

$$\underline{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_h], \quad (3.5)$$

the minimum of J is given by

$$\frac{\partial J}{\partial \gamma_j} = 0, \quad j = 1, 2, \dots, h. \quad (3.6)$$

This expression provides h equations for the h unknown parameters.

4. SINGLE UNKNOWN PARAMETER IN A SINGLE FIRST-ORDER SYSTEM

The parameter estimation process is introduced by considering a simple system described by a single first-order equation involving a single unknown parameter. The system plant or state equation is assumed to be given by

$$\dot{x}(t) = a x(t) + u(t), \quad (4.1)$$

where $x(t)$ is the state variable, $u(t)$ is the input variable and a is the unknown parameter which is assumed constant. This equation takes the form of Equation (3.1a) in which the matrix $\underline{A}$ has been reduced to the scalar quantity a and the matrix $\underline{B}$ has been reduced to unity.

The object of the estimation process is to determine the value of the parameter a in Equation (4.1) from measurements of the plant output and plant input at the discrete times $t = t_i$, $i = 1, 2, \dots, N$. The particular output variable available depends on circumstances and the type of system under consideration. When only those variables that are directly present in Equation (4.1) are to be used, three possible combinations can be identified:

- (i) when measurements of the state variable, $x(t)$, and the input variable, $u(t)$, are available,
- (ii) when measurements of the rate of change of the state variable with respect to time, $\dot{x}(t)$, and the input variable, $u(t)$, are available, and
- (iii) when measurements of $\dot{x}(t)$, $x(t)$ and $u(t)$ are all available.

A parameter estimation process is described for each of these three cases in Sections 4.1, 4.2 and 4.3, respectively. For Case (i) and Case (ii) two classes of technique may be identified. The first treats the state equation as essentially an algebraic relationship between $\dot{x}(t)$, $x(t)$, and $u(t)$ and leads to a linear equation in the unknown parameter, a . The method consequently provides a direct solution for the unknown parameter. The second class of method solves the differential equation, Equation (4.1), in terms of the input quantity which leads to a nonlinear equation in the parameter a . This method now involves an iterative solution for the unknown parameter.

To illustrate the application of the methods a simple example has been included in each of the following sections.

4.1 Measurement of State Variable and Input Variable

In this section it is assumed that the state variable, $x(t)$, and the input variable, $u(t)$, are measured and denoted by,

$$z_x(t) = \text{the measured values of } x(t) \quad (4.2a)$$

$$\text{and } z_u(t) = \text{the measured values of } u(t). \quad (4.2b)$$

The difference between the measured and predicted output may be expressed as

$$\varepsilon_x(t) = z_x(t) - x_s(t), \quad (4.3)$$

where the quantity $x_s(t)$ in Equation (4.3) is a solution of $x(t)$ found from the state equation of the system in terms of the measured quantities $z_x(t)$ and $z_u(t)$, and the unknown parameter, a .

This page Amendment B

4.1.1 Direct solution for unknown parameter

The state equation can be rearranged to an integral form that is linear in the parameter, a , given by

$$x(t) = a \int_0^t x(t') dt' + \int_0^t u(t') dt'. \quad (4.4)$$

Thus one representation of $x_s(t)$, *i.e.* a representation of $x(t)$ in terms of measurements, is

$$x_s(t) = a \int_0^t z_x(t') dt' + \int_0^t z_u(t') dt', \quad (4.5)$$

or
$$x_s(t) = a \tilde{z}_x(t) + \tilde{z}_u(t), \quad (4.6)$$

where the curly bar above the measurement quantities indicates the integral of that quantity between the limits 0 and t .

The difference function, $\epsilon_x(t)$, given by Equation (4.3), may now be rewritten, using Equation (4.6), as

$$\epsilon_x(t) = z_x(t) - a \tilde{z}_x(t) - \tilde{z}_u(t). \quad (4.7)$$

When measurements are taken at the discrete times, t_i , for $i = 1, 2, \dots, N$, then Equation (4.7) is rewritten as

$$\epsilon_{xi} = z_{xi} - a \tilde{z}_{xi} - \tilde{z}_{ui}, \quad (4.8)$$

where

$$\epsilon_{xi} = \epsilon_x(t_i), \quad (4.9a)$$

$$z_{xi} = z_x(t_i), \quad (4.9b)$$

$$\tilde{z}_{xi} = \begin{cases} 0 & , i = 1 \\ \sum_{j=1}^{i-1} 0.5(z_{xj} + z_{x(j+1)})(t_{j+1} - t_j) & , i > 1 \end{cases} \quad (4.9c)$$

and

$$\tilde{z}_{ui} = \begin{cases} 0 & , i = 1 \\ \sum_{j=1}^{i-1} 0.5(z_{uj} + z_{u(j+1)})(t_{j+1} - t_j) & , i > 1 \end{cases}. \quad (4.9d)$$

The Equations (4.9c) and (4.9d) have used trapezoidal integration which is probably adequate bearing in mind that the measurement values contain noise.

The Objective Function, J , to be minimised takes the form

$$J = \sum_{i=1}^N \epsilon_{xi}^2, \quad (4.10)$$

which may be rewritten using Equation (4.8) as,

$$J = \sum_{i=1}^N \{z_{xi} - a \tilde{z}_{xi} - \tilde{z}_{ui}\}^2. \quad (4.11)$$

This equation is a minimum when $\partial J / \partial a = 0$, i.e. when

$$\sum_{i=1}^N \{z_{xi} - a\tilde{z}_{xi} - \tilde{z}_{ui}\}\tilde{z}_{xi} = 0, \quad (4.12)$$

which is rearranged to give

$$a = \left\{ \sum_{i=1}^N \tilde{z}_{xi}^2 \right\}^{-1} \left\{ \sum_{i=1}^N z_{xi}\tilde{z}_{xi} - \sum_{i=1}^N \tilde{z}_{ui}\tilde{z}_{xi} \right\}. \quad (4.13)$$

It is usual to denote the unknown parameter obtained from Equation (4.13) by the symbol $\hat{a}$ since it is the least-squares estimate of the parameter, a .

4.1.2 Iterative solution for unknown parameter (Output Error method)

The method described here, known in the literature as the Output Error method, uses an alternative solution, $x_s(t)$, to the state equation in terms of the measured input quantities, which may be defined as a solution of

$$\dot{x}_s(t) = a x_s(t) + z_u(t). \quad (4.14)$$

A solution of this equation for a given value of a is denoted by $x_s(a, t)$ and may be obtained analytically[†] from

$$x_s(a, t) = \int_0^t e^{a(t-\tau)} u(\tau) d\tau, \quad (4.15)$$

or alternatively by a numerical integration method, such as Runge-Kutta.

Since this equation is nonlinear in the unknown parameter, a technique best described as linearisation is used where the increment in the unknown parameter, Δa , is related to a new estimate for a , by

$$a' = a_I + \Delta a, \quad (4.16)$$

where a_I is an initial estimate for a . Using a Taylor series expansion of $x_s(a, t)$ about the initial estimate of the unknown parameter, the following approximation is obtained

$$x_s(a, t) = x_s(a_I, t) + \frac{\partial x_s(a_I, t)}{\partial a} \Delta a + \text{higher-order terms}, \quad (4.17)$$

where $x_s(a_I, t)$ is the solution of the state equation with $a = a_I$.

[†] It may be noticed that the state equation is a standard form of differential equation usually described in mathematical texts as

$$\frac{dx}{dt} + P(t)x = Q(t),$$

which may be solved by the "Integrating Factor" method.

In discrete form, retaining only first-order terms, this equation is

$$x_s(a)_i = \left(x_s(a_I)\right)_i + \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i \Delta a, \quad (4.18)$$

where i is an integer denoting the individual values of the independent variable t .

The difference between the measured and predicted output is given by Equation (4.3), which in discrete form becomes

$$\epsilon_{xi} = z_{xi} - x_{si}. \quad (4.19)$$

By use of Equation (4.19), the Objective Function, J , to be minimised, takes the form of Equation (4.10).

$$J = \sum_{i=1}^N \epsilon_{xi}^2 = \sum_{i=1}^N (z_{xi} - x_{si})^2, \quad (4.20)$$

which may be rewritten using Equation (4.18) as

$$J = \sum_{i=1}^N \left\{ z_{xi} - \left(x_s(a_I)\right)_i - \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i \Delta a \right\}^2. \quad (4.21)$$

To obtain the partial derivative, $\partial x_s(a_I, t)/\partial a$, representing the sensitivity of the model to changes in parameter values, the method of differential corrections is usually used which is more accurate and considerably faster than the perturbation method. The parametric differentiation technique involves differentiating the state equation with respect to the parameter, which gives

$$\frac{d}{dt} \left(\frac{\partial x_s(a_I, t)}{\partial a} \right) = a_I \frac{\partial x_s(a_I, t)}{\partial a} + x_s(a_I, t), \quad \frac{\partial x_s(a_I, 0)}{\partial a} = 0, \quad (4.22)$$

in which the order of differentiation has been reversed. The Objective Function, J , from Equation (4.21) is a minimum when $\partial J(a)/\partial a = 0$, i.e. when

$$\sum_{i=1}^N \left\{ z_{xi} - \left(x_s(a_I)\right)_i - \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i \Delta a \right\} \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i = 0, \quad (4.23)$$

$$\text{or} \quad \Delta a = \frac{\sum_{i=1}^N z_{xi} \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i - \sum_{i=1}^N \left(x_s(a_I)\right)_i \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i}{\sum_{i=1}^N \left(\frac{\partial x_s(a_I)}{\partial a}\right)_i^2}. \quad (4.24)$$

This equation provides the increment Δa which may be used in Equation (4.16) to obtain a new estimate of the unknown parameter, a' . With this new estimate the complete process can be repeated until a minimum of the cost function is obtained, *i.e.* when Δa is sufficiently small.

The linearisation procedure used above is the first-order or Gradient method of Newton-Raphson. A second-order optimisation procedure, a Modified Newton-Raphson procedure, commonly called Newton-Gauss, will be additionally described.

In this method the Taylor series expansion retains terms up to and including those of second-order. To obtain an expression for Δa it is simpler to expand the Objective Function in a Taylor series directly about an initial value of the unknown parameter, rather than expanding $x_s(a, t)$ as shown by Equation (4.17). Expanding $J(a)$ gives

$$J(a) = J(a_I) + (a - a_I) \frac{\partial J(a_I)}{\partial a} + \frac{(a - a_I)^2}{2} \frac{\partial^2 J(a_I)}{\partial a^2} + \text{third-order terms.} \quad (4.25)$$

Differentiating this equation with respect to a and discarding third- and higher-order terms, the following expression is obtained

$$\frac{\partial J(a)}{\partial a} = 2 \frac{\partial J(a_I)}{\partial a} + (a - a_I) \frac{\partial^2 J(a_I)}{\partial a^2}. \quad (4.26)$$

This equation is equal to zero for $J(a)$ to be a minimum, which gives

$$a = a_I - \left\{ \frac{\partial^2 J(a_I)}{\partial a^2} \right\}^{-1} \left\{ \frac{\partial J(a_I)}{\partial a} \right\}, \quad (4.27)$$

or, in terms of the increment in the parameter,

$$\Delta a = - \left\{ \frac{\partial^2 J(a_I)}{\partial a^2} \right\}^{-1} \left\{ \frac{\partial J(a_I)}{\partial a} \right\}. \quad (4.28)$$

The first and second derivatives of the Objective Function may be obtained from Equation (4.20), as

$$\frac{\partial J}{\partial a} = 2 \sum_{i=1}^N \epsilon_{xi} \frac{\partial \epsilon_{xi}}{\partial a} \quad (4.29)$$

and

$$\frac{\partial^2 J}{\partial a^2} = 2 \sum_{i=1}^N \left(\frac{\partial \epsilon_{xi}}{\partial a} \right)^2 + 2 \sum_{i=1}^N \epsilon_{xi} \frac{\partial^2 \epsilon_{xi}}{\partial a^2}. \quad (4.30)$$

Usually the second term of Equation (4.30) is dropped and the algorithm in this form is then known as the Gauss-Newton approximation. Near the solution the neglected term is generally small and the Gauss-Newton expression approaches the excellent performance of the Newton-Raphson method. It is noted that the second derivative of the Objective Function need never be directly calculated since it contains only the first derivative of the residual error.