General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

UNH-

N67 12987
(ACCESSION NUMBER)

(THRU)

11
(PAGES)

1
(CODE)

CR 80355
(NASA CR OR TMX OR AD NUMBER)

19
(CATEGORY)

FACILITY FORM 602

GPO PRICE $ ________

CFSTI PRICE(S) $ ________

Hard copy (HC) 1.00

Microfiche (MF) .50

ff 653 July 65

THE UNIVERSITY OF NEW HAMPSHIRE · 1923

Department of Physics

UNIVERSITY OF NEW HAMPSHIRE

Durham

MAXIMUM LIKELIHOOD METHOD FOR FITTING A SUM OF EXPONENTIALS TO EXPERIMENTAL DATA

By R.A. Beauregard

University of New Hampshire
Department of Physics
Durham, New Hampshire

This work was supported by the National Aeronautics and Space Administration under Grant #NGR-30-002-018 and Grant #NGR-30-002-021, and by the University of New Hampshire Computation Center Graduate Assistantship.

MAXIMUM LIKELIHOOD METHOD FOR FITTING A SUM OF EXPONENTIALS TO EXPERIMENTAL DATA

R. A. Beauregard

INTRODUCTION

This paper describes a procedure for fitting a sum of exponentials to a set of experimental data points by the method of maximum likelihood as described by Grard [1]. In particular, an iterative process is used to maximize the likelihood function. Part I describes how the likelihood function may be written when a normal distribution of deviations in the observed data is assumed. A brief description of the error matrix is given. Part II describes a modified computer program (Fortran) designed for use at the University of New Hampshire using the IBM 1620 and the IBM 360. A sample set of data to be fit by two exponentials is given along with the solution by the program described.

---

[1] The computer program is a modified version of the "Malik" program by Grard, Lawrence Radiation Laboratory, University of California (UCRL-10153). See also, "A general Program for Statistical Analysis Using the Maximum-Likelihood Method (Malik) Program"; F. Grard Nucl. Inst. and MTD's, 34, 242-244 (1965).

PART I

MAXIMUM LIKELIHOOD APPROACH

Assume n measurements $y_1 .., y_i .., y_n$. are made at times $t_1 .. t_i .. t_n$ in some experiment, and it is desired to fit this data to a sum of exponentials given by:

$$f(t) = y = \sum_{j=1}^{K} a_j e^{-b_j t} + a_{k+1}$$

where the a_j and b_j are parameters to be determined. The problem then, is to find values of a_j ($j=1, \cdots k+1$) amd b_j ($j=1, \ldots K$) that gives the best fit of this function to the data. Further, assume that the deviations of the y_i from $y=f(t_i)$ are normally distributed. The deviations are given by:

$$\Delta y_i = y_i - f(t_i)$$

Then the probability of observing a value of y_i in the interval $(y_i, y_i + dy_i)$ is given by:

$$p(y_i + dy_i) = \frac{1}{\sqrt{2\pi}} \frac{dy_i}{\sigma_i} \exp\left[- \frac{(\Delta y_i)^2}{2(\sigma_i)^2} \right] \qquad (1)$$

where $(\sigma_i)^2$ is the weighting factor to be associated with y_i. The likelihood function is the joint probability of observing simultaneously the measurements $y_1 .., y_i .., y_n$ in the intervals $(\Delta y_1 \ldots \Delta y_i \ldots \Delta y_n)$ and is given by:

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \frac{dy_i}{\sigma_i} \exp\left[- \frac{(\Delta y_i)^2}{2(\sigma_i)^2} \right], \qquad (2)$$

This assumes all y_i measurements are independent. L in (2) may be considered as a function of the parameters a_j and b_j (which appear in Δy_i) and we seek values of these parameters such that L is maximum. To maximize L, it is convenient (and sufficient) to maximize log L. Hence we define the function F by

$$F = \log_e L = \sum_{i=1}^{n} \left[\log_e \left(\frac{1}{\sqrt{2\pi}} \frac{dy_i}{\sigma_i} \right) - \frac{(\Delta y_i)^2}{2(\sigma_i)^2} \right] \quad (3)$$

It is evident that maximizing F is equivalent to maximizing $\bar{F}$ * given by:

$$\bar{F} = -\frac{1}{2} \sum_{i=1}^{n} \left(\frac{\Delta y_i}{\sigma_i} \right)^2 \quad (4)$$

The values of the parameters that maximize equation (4) are the values sought to fit (1) to the observed data. One method that may be used to maximize (4) is described in Part II.

The error matrix associated with the best values of the parameters may be expressed as follows. For notational convenience, we re-label the parameters as follows:

$$\begin{array}{ll} \phi_1 = a_1 & \phi_2 = b_1 \\ \phi_3 = a_2 & \phi_4 = b_2 \\ \vdots & \vdots \end{array}$$

---

* This is evidently the "Least Squares" criterion. Note the assumptions that have been made concerning the deviations in the measured data in this particular application of the method of maximum likelihood; that is, the deviations are normally distributed.

Let $\hat{\phi}_1 \ldots \hat{\phi}_{2K+1}$ denote the best values of $\phi_1 \ldots \phi_{2K+1}$ respectively. We assume that $L(\phi_1 \ldots \phi_{2K+1})$ is a normal (multi-gaussian) distribution. Then we may write:

$$L(\phi_1, \ldots \phi_{2K+1}) = K \exp\left[-\frac{1}{2} \sum_{i,j=1}^{2K+1} A_{ij}(\phi_i - \hat{\phi}_i)(\phi_i - \hat{\phi}_j)\right] \quad (5)$$

where the Matrix $||A_{ij}||$ is symmetric, and where the variance-co-variance of ϕ_i and ϕ_j is given by:

$$||\sigma_i \sigma_j \rho_{ij}|| = ||A_{ij}||^{-1}$$

To find a convenient expression for the A_{ij}, we proceed as follows. The log of equation (5) yields:

$$F = \text{Log}L = \text{Log}K - \frac{1}{2} \sum_{i=1}^{2K+1} \left(\sum_{j=1}^{2K+1} \left(A_{ij}(\phi_i \phi_j - \hat{\phi}_i \phi_j - \hat{\phi}_j \phi_i + \hat{\phi}_i \hat{\phi}_j) \right) \right)$$

We now calculate all 2nd order partials of F.

$$\frac{\delta F}{\delta \phi_\tau} = -\frac{1}{2} \sum_{i=1}^{2K+1} \sum_{j=1}^{2K+1} \frac{\delta}{\delta \phi_\tau} [A_{ij}(\phi_i \phi_j - \hat{\phi}_i \phi_j - \hat{\phi}_j \phi_i + \hat{\phi}_i \hat{\phi}_j)]$$

$$= -\frac{1}{2} \sum_{\substack{i=1 \\ i \neq j}}^{2K+1} \frac{\delta}{\delta \phi_\tau} [A_{ij}(\phi_i \phi_i \quad \hat{\phi}_i \phi_i \quad \hat{\phi}_i \phi_i \quad \hat{\phi}_i \hat{\phi}_i)]$$

$$- \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{2K+1} \frac{\delta}{\sigma \phi_\tau} [A_{ij}(\phi_i \phi_j - \hat{\phi}_i \phi_j - \hat{\phi}_j \phi_i + \hat{\phi}_i \hat{\phi}_j)]$$

$$- \frac{1}{2} \frac{\delta}{\delta \phi_\tau} [A_{ij}(\phi_i \phi_i - \hat{\phi}_i \phi_i - \hat{\phi}_i \phi_i + \hat{\phi}_i \hat{\phi}_i)]$$

where the terms not appearing are zero. Thus:

$$\frac{\delta F}{\delta \phi_\tau} = -\frac{1}{2} \sum_{\substack{i=1 \\ i \neq \tau}}^{2K+1} [A_{i\tau}(\phi_i - \hat{\phi}_i)] - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq t}}^{2K+1} [A_{\tau j}(\phi_j - \hat{\phi}_j)] - \frac{1}{2} A_{\tau\tau}(2\phi_\tau - 2\hat{\phi}_\tau)$$

$$\frac{\delta^2 F}{\delta \phi_s \, \delta \phi_\tau} = \begin{cases} -\frac{1}{2} A_{s\tau} - \frac{1}{2} A_{s\tau}, & s \neq \tau \\ (-\frac{1}{2} A_{\tau\tau}) \quad (2), & s = \tau \end{cases}$$

$$\frac{\delta^2 F}{\delta \phi_s \delta \phi_\tau} = - A_{s\,\tau} \qquad (6)$$

Equation (6) gives a convenient form of the error matrix. That is, the variance-co-variance matrix may be calculated with all 2nd order derivatives of the log of the likelihood function. Therefore:

$$||\sigma_i \sigma_j P_{ij}|| = ||A_{ij}||^{-1}$$

$$= || - \frac{\delta^2 F}{\delta \phi_i \delta \phi_j} ||^{-1}$$

PART II

COMPUTER PROGRAM

The program described in this paper has been written for the analysis of experimental data by use of the method of maximum likelihood. The main program maximizes the likelihood function by a simple iterative procedure, and also calculates the error matrix. Three subroutines are used:

a) Matinv to invert a matrix

b) Funct to calculate F=logL

c) Input to read and store the observed data

The latter two subroutines must be written by the user only if he wishes to solve problem other than the one described in Part I. All programs are written in Fortran II.

<u>Iterative Process in Main Program</u>

The method used to maximize F, the log of the likelihood function, is the following. First initial guesses for the parameters a_j, b_j are made and F is calculated. The first parameter is increased or decreased by a quantity corresponding to its assigned step in order to find a larger value of F. If a larger value of F is found, the new value of this parameter is used and the same operation is repeated for the next parameter, etc. This cycle is repeated as long as F increases. If no improvement can be made with the step sizes being used, the steps are all reduced by a common factor a, and the next iteration is started. The iterative process continues until the difference between the max values of F in the j^{th} and $(j-1)^{th}$ iterations respectively is less than 0.001.

In a given iteration, if a parameter does not yield a larger value of $\bar{F}$ after three consecutive attempts, its step is reduced by the factor a and the same iteration continues. In this case, this step will not be reduced for the next iteration.

Error Matrix in Main Program

The error matrix is computed by inversion of the matrix built with all second derivatives of $\bar{F}$ at its maximum. The second derivatives are calculated by differences. In order to find an adequate Δ to be used, the width of the likelihood function L in the direction of each parameter is determined as being a positive and a negative quantity one must add to each of the fitted parameters to Lower L by $e^{-1/2}$ of its maximum value (i.e. to lower F by $-\frac{1}{2}$) The Δ's used to compute the second differences are taken to be a fraction B of the width of L. B must be read in as initial information. The second derivative matrix is the pointwise mean of the two secon difference matrices. The error matrix is the inverse of the second derivative matrix.

During the calculation of the second differences, it is possible that a larger value of $\bar{F}$ is obtained, if the iteration had stopped at a saddle point. In this case the iterative process is re-started with the initial steps.

Input Subroutine

In general, the input subroutine is used once to read the observed data points y_i, t_i, σ_i. These values are used exclusively by the Funct subroutine.

Funct Subroutine

In general, the Funct subroutine is used repeatedly to calculate F = log L in terms of the parameters and the observed data points.

The Funct subroutine that was written for the problem in Part I calculates:

$$F = -\frac{1}{2} \sum_{j=1}^{n} \left[\frac{y_j - \sum_{i=1}^{K} a_i e^{-b_i t_j}}{\sigma_j} \right]^2$$

Input Data

The data cards read by the main program are the following:

1) One card with the number of parameters to be fitted.
Format (20 x, I10) This number must not exceed 10.

2) At most ten data cards, one for each parameter (initial guess) and the corresponding assigned step.
Format (10 x, 2F10.5)

3) One card containing the factor a used to reduce the steps at the end of each iteration, the total number of iterations allowed, and the factor B used to calculate second differences.
Format (20 x, F10.5, I10, F10.5)

The main program then calls the Input subroutine. The data cards read by the Input subroutine that was written for the Problem in Part I are the following.

4) One card containing the number of data points (maximum is 40). Format (I5).

5) At most 40 cards, each containing the values t_i, y_i, σ_i Format (F10.5, 2F10.2)

SAMPLE SET OF DATA FITTED TO THE SUM OF TWO EXPONENTIALS

$$y_i = a_1 e^{-b_1 t_i} + a_2 e^{-b_2 t_i}$$

y_i	t_i	σ_i
26268	.0204	271.94
14636	.1105	168.73
11200	.1561	145.74
8942	.2008	109.49
6465	.2464	96.28
5308	.2913	74.81
4211	.3362	70.19
3171	.3814	62.61
2543	.4264	59.05
2166	.4714	56.37
1718	.5173	51.71
1561	.5622	41.83
1266	.6069	39.91
1172	.6519	38.87
952	.6971	37.40
903	.7425	33.35
747	.7874	32.38
793	.8328	32.85
618	.8778	31.71
671	.9321	32.99
492	.9682	32.09
607	1.0134	34.65
579	1.0584	37.33

$a_1 = 28320 \pm 2.62$

$b_1 = 6.519 \pm 0.042$

$a_2 = 1245 \pm 101$

$b_2 = 0.86 \pm 0.08$