UNCLASSIFIED

ORNL-455
Series B
Physics-General

ON PRONY'S METHOD OF FITTING EXPONENTIAL DECAY CURVES AND MULTIPLE-HIT SURVIVAL CURVES

OAK RIDGE NATIONAL LABORATORY
OPERATED BY
CARBIDE AND CARBON CHEMICALS DIVISION
UNION CARBIDE AND CARBON CORPORATION
UCC
POST OFFICE BOX P
OAK RIDGE, TENNESSEE

UNCLASSIFIED

UNCLASSIFIED

ORNL 455, Series B

This document consists of
16 pages.
Copy 14 of 15 Series B.

Contract No. W-7405, eng 26

MATHEMATICS AND COMPUTING PANEL

ON PRONY'S METHOD OF FITTING EXPONENTIAL DECAY CURVES AND MULTIPLE-HIT SURVIVAL CURVES

A. S. Householder

DATE ISSUED: FEB 3 - 1950

OAK RIDGE NATIONAL LABORATORY
operated by
CARBIDE AND CARBON CHEMICALS DIVISION
Union Carbide and Carbon Corporation
Post Office Box P
Oak Ridge, Tenn.

UNCLASSIFIED

2

ORNL 455
Physics-General

DISTRIBUTION FOR SERIES B:

1-15. Technical Information Branch, ORE

ON PRONY'S METHOD OF FITTING EXPONENTIAL DECAY CURVES AND MULTIPLE-HIT SURVIVAL CURVES

Prony's method of fitting exponentials is described by Whittaker and Robinson (3), and by Willers (4). The method is extremely elegant but suffers from two serious drawbacks. First, while having the appearance of a least-squares curve fitting, it is not strictly such, and in particular it provides no means for weighting the observations in accordance with their supposed precision. Second, it provides no criterion for determining the number of exponentials required for the fitting, so that one must either suppose this known in advance, or else, perhaps, apply the method repeatedly, increasing the number of terms until one is satisfied with the result. Each time this is done the work must start afresh.

It is quite possible, however, to elaborate Prony's method somewhat so as to obtain a valid least squares fit with correct weighting of the measurements. It is also possible to develop a sequential process such that, say, at the n-th step one can test whether n+1 exponentials are required, after which one either proceeds to step n+1, or utilizes the computed results in a simple fashion to complete the first step in the calculation. For obtaining a valid least squares fit it is necessary to utilize the results of the first step in Prony's method as a first approximation, to be improved by an appropriate iteration. We shall therefore first describe Prony's method, which seems to be not well known. Thereafter it will be shown how one can proceed to obtain a valid least squares fit, and finally we shall indicate the criterion for choosing the number of exponentials required.

The method is adapted to cases where measurements are made at equally spaced intervals of time. This is not the usual way in which, for example, radio-active decay data is presented, but in many instances an obvious interpolation in the records will provide this. Hence, if the theoretical function is

$$y(t) = \sum_{i=1}^{n} A_i \exp(-\alpha_i t), \tag{1}$$

and measurements are made at times

$$t = 0, \tau, 2\tau, \ldots, N\tau$$

we may set

$$u_i = \exp(-\alpha_i \tau); \tag{2}$$

then the experimental values $Y_0, Y_1, \ldots, Y_N$ are estimates of the quantities

$$y_r = y(r\tau) = \sum_{i=1}^{n} A_i u_i^{\,r}. \tag{3}$$

The method is based upon the fact that the y_r must satisfy a difference equation of order n which may be written in the form

$$y_p s_n + y_{p+1} s_{n-1} + \ldots + y_{p+n-1} s_1 + y_{p+n} = 0, \tag{4}$$

the s_i being, apart from sign, the elementary symmetric functions of the u_i. If the s_i can be determined, then we have an equation of degree n to be solved for the u_i, and the A_i can be obtained thereafter by ordinary least squares.

Prony's method continues now to form the normal equations from (4) after the usual pattern, obtaining the set

$$[Y_0\ Y_0]\, s_n + [Y_0\ Y_1]\, s_{n-1} + \ldots + [Y_0\ Y_n] = 0,$$

$$[Y_1\ Y_0]\, s_n + [Y_1\ Y_1]\, s_{n-1} + \ldots + [Y_1\ Y_n] = 0, \qquad (5)$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$[Y_{n-1}\ Y_0]\, s_n + [Y_{n-1}\ Y_1]\, s_{n-1} + \ldots + [Y_{n-1}\ Y_n] = 0,$$

where

$$[Y_p\ Y_q] = Y_p\, Y_q + Y_{p+1}\, Y_{q+1} + \ldots \qquad (6)$$

It is clear that these equations are formed in accordance with the usual least-squares model. However, the usual assumptions do not hold in this case, since the coefficients in these equations are subject to error, whereas in the ordinary derivation of the normal equations it is supposed that only the terms free of the unknowns (in this case the s_i) are subject to error. Furthermore, as we remarked above, there is no way provided in this method for applying statistical weights to the Y's when they differ in precision.

Since the method purports to yield a least-squares fit to equations (4), it is at once suggested that we examine the correct method for obtaining this. The correct method is to so determine the N+1 quantities y_r and the n parameters s_i that the sum of squares

$$S = \Sigma\, w_r\, (Y_r - y_r)^2 \qquad (7)$$

is minimized subject to the fulfillment of the side conditions (4).

The multipliers w_r are the statistical weights associated with the measurements $\bar{Y}_r$. If we introduce the Lagrange multipliers λ_p and define

$$s = \sum_{r=1}^{N} w_r (\bar{Y}_r - y_r)^2 - 2 \sum_{p}^{N-n} \lambda_p P_p \tag{8}$$

where

$$\bar{P}_p \equiv y_p s_n + y_{p+1} s_{n-1} + \ldots + y_{p+n} = 0, \tag{9}$$

then the minimum is obtained by adjoining to equations (4) the equations

$$\partial s / \partial y_r = 0, \qquad \partial s / \partial s_i = 0 \tag{10}$$

and solving for the unknowns s_i, λ_p, and y_r. Unfortunately, the equations are non-linear and their solution does not appear to be easy.

We may, however, proceed as follows: Presumably the solution provided by Prony's method would give at least a first approximation to the theoretically correct solution. Consequently, we may accept this as an approximation and proceed thereafter to improve sequentially the results. This is done by expanding in Taylor's series, retaining only constant and linear terms, and solving [see Deming (2), Czuber (1)]. For this we set

$$\bar{Y}_r = y_r + \eta_r, \quad r_i = s_i + \rho_i, \tag{11}$$

where the r_i are solutions of equations (5), the ρ_i and the η_r are (presumably small) corrections. We can calculate

$$\bar{P}_{p0} = \bar{Y}_p r_n + \bar{Y}_{p+1} r_{n-1} + \ldots + Y_{p+n}, \tag{12}$$

and by neglecting terms of second order in the ρ_i and η_r write, in place of (9),

$$\bar{P}_{p0} = (\eta_p r_n + \ldots + \eta_{p+n}) + (\bar{Y}_p \rho_n + \ldots + \bar{Y}_{p+n-1}\rho_1). \tag{13}$$

If these linearized equations are used in place of (9) the problem reduces to that of solving a system of linear equations for the unknowns η, ρ and λ. We proceed as follows:

Denote by λ the column vector of the λ's, by ρ the column vector of the ρ's in reversed order, by η the column vector of the η's, and by P the column vector of the P_0's. Let W designate the diagonal matrix of the w's. Form the matrices

$$R = \begin{pmatrix} r_n & 0 & 0 & \cdots & 0 \\ r_{n-1} & r_n & 0 & \cdots & 0 \\ r_{n-2} & r_{n-1} & r_n & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_1 & r_2 & r_3 & \cdots & \cdots \\ 1 & r_1 & r_2 & \cdots & \cdots \\ 0 & 1 & r_1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{14}$$

and

$$\bar{Y} = \begin{pmatrix} \bar{Y}_0 & Y_1 & \cdots & Y_{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ Y_{N-n} & \bar{Y}_{N-n+1} & \cdots & \bar{Y}_n \end{pmatrix} . \quad (15)$$

Then the equations to be solved are found to be

$$\begin{aligned} W\eta - R\lambda &= 0, \\ R'\eta + Y\rho &= \bar{P} \\ Y'\lambda &= 0 . \end{aligned} \quad (16)$$

Hence the first of these matrix equations gives

$$\eta = W^{-1} R\lambda,$$

and then the second becomes

$$R' W^{-1} R\lambda = \bar{P} - Y\rho .$$

If we set

$$Q = R' W^{-1} R,$$

we have

$$\lambda = Q^{-1} \bar{P} - Q^{-1} \bar{Y}\rho .$$

Hence from the third,

$$Y' Q^{-1} Y\rho = \bar{Y}' Q^{-1} P.$$

If we solve these equations for ρ, we are able then to find first λ and then η by substituting back into the preceding equations. It is to be noted that the matrix Q is a triangular matrix and therefore easy to invert; the last equation in λ involves only λ_{N-n}, the preceding one involves only λ_{N-n}, and λ_{N-n-1}, etc. The equations can therefore be solved sequentially starting with the last and proceeding toward the first.

When the solution is complete, if the η's and ρ's turn out to be fairly large it may be necessary to repeat. One substracts the η's thus found from the Y's, the ρ's from the r's, renames the results Y and r, and proceeds exactly as before.

The statistical test for goodness of fit consists in computing

$$s = (\check{Y} - y)' \quad W(\check{Y} - y), \tag{17}$$

where Y is the vector of initial measurements and y the finally accepted adjusted values. This has a chi-square distribution with N-n degrees of freedom.

After the s_i are determined one must solve the algebraic equation

$$u^n + s_1 u^{n-1} + \ldots + s_n = 0, \tag{18}$$

and the α's are obtained from these n roots u_i by

$$\alpha_i = -\tau^{-1} \log u_i.$$

In using Prony's method one must replace equations (3) by the normal equations with Y in place of y, and solve for the A's by least squares. However if the adjusted values y_r are used, and these have been obtained with sufficient accuracy, then any set of n of the equations (3) will be suitable, and the normal equations are not necessary unless greater accuracy is required.

Our modified Prony's method is seen to consist of four distinct steps. The first is to form the pseudo-normal equations (5) and solve for the s_i. These values are not, however, the true least squares estimates we require, so we rename them r_i and take the second step which amounts to solving equations (16) for the ρ's and η's, using them to correct the r's and adjust the Y's according to (11). The third step consists in solving the algebraic equation (18) whose coefficients are the quantities s_i just found. The last step consists in solving any set of n equations (3), or the normal equations obtained from the entire set, for the coefficients A_i. But we have, as yet, no criterion for determining the number n of exponentials required. We return, therefore, to the first step involving the pseudo-normal equations.

Consider the vectors

$$v_0 = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_\gamma \end{pmatrix}, \; v_1 = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ y_{\gamma+1} \end{pmatrix}, \ldots, v_i = \begin{pmatrix} y_i \\ y_{i+1} \\ y_{i+2} \\ \cdot \\ \cdot \\ \cdot \\ y_{\gamma+i} \end{pmatrix}, \ldots, \tag{19}$$

where, for the moment, we leave γ unspecified. If the Y_r satisfied a difference equation of order n, then so would the vectors v_i, and, in fact, any n+1 of these vectors would be linearly dependent. However, if $\gamma \geq n$, any smaller number of these vectors would be linearly independent.

We could write, in fact,

$$v_0 s_n + v_1 s_{n-1} + \dots + v_n = 0. \tag{20}$$

Let us apply Choleski's orthogonalization process to these equations.

This process consists of first orthogonalizing the vectors v_i. That is, we replace each vector v_i by a linear combination of the vectors $v_0, v_1, \dots, v_i$ orthogonal to all vectors $v_0, v_1, \dots, v_{i-1}$. Thus we set

$$v_0 = a_0,$$

$$v_1 = a_1 + \mu_{10} a_0,$$

where μ_{10} is chosen so as to make a_1 and a_0 orthogonal. This means that

$$a_0' v_1 = \mu_{10} a_0' a_0 ;$$

Next, we set

$$v_2 = a_2 + \mu_{21} a_1 + \mu_{20} a_0 ,$$

choosing μ_{21} and μ_{20} so that a_2 is orthogonal to both a_0 and a_1. Hence

$$a_0' v_2 = \mu_{20} a_0' a_0,$$

$$a_1' v_2 = \mu_{21} a_1' a_1.$$

Proceeding sequentially, we finally set

$$v_{n-1} = a_{n-1} + \mu_{n-1,n-2} a_{n-2} + \dots + \mu_{n-1,0} a_0$$

where

$$a'_{n-2} \quad v_{n-1} \;=\; \mu_{n-1,n-2} \; a'_{n-2} \; a_{n-2} \;,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a'_0 \quad v_{n-1} \;=\; \mu_{n-1,0} \; a'_0 \; a_0 \;.$$

Let V_i, A_i, M_i designate the matrices

$$V_i = (v_0, v_1, \ldots, v_i), \tag{21}$$

$$A_i = (a_0, a_1, \ldots, a_i), \tag{22}$$

$$M_i = \begin{pmatrix} 1 & \mu_{10} & \mu_{20} & \cdots & \mu_{i0} \\ 0 & 1 & \mu_{21} & \cdots & \mu_{i1} \\ 0 & 0 & 1 & \cdots & \mu_{i2} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \end{pmatrix} . \tag{23}$$

Then

$$V_i = A_i \, M_i \quad . \tag{24}$$

Now M_i, being a triangular matrix, is easily inverted, so that the solution

$$A_i = V_i \, M_0^{-1} \tag{25}$$

is easily obtained.

Equations (20) can be written in the form

$$V_{n-1}\,\sigma_n + v_n = 0, \tag{26}$$

if by σ_n we designate the column vector

$$\sigma_n = \begin{pmatrix} s_n \\ s_{n-1} \\ \vdots \\ \cdots \\ \cdots \\ s_1 \end{pmatrix}$$

By means of (24) we see that (26) is equivalent to

$$A_{n-1}\,M_{n-1}\,\sigma_n + v_n = 0. \tag{27}$$

But since A_n is a matrix of orthogonal vectors, the product

$$D_{n-1} = A'_{n-1}\,A_{n-1} \tag{28}$$

is a diagonal matrix. Consequently (27) can be written

$$D_{n-1}\,M_{n-1}\,\sigma_n + A'_{n-1}\,v_n = 0, \tag{29}$$

and this gives

$$M_{n-1}\,\sigma_n + D^{-1}_{n-1}\,A'_{n-1}\,v_n = 0, \tag{30}$$

by an easy inversion of D_{n-1}, and thence

$$\sigma_n + M^{-1}_{n-1}\,D^{-1}_{n-1}\,A'_{n-1}\,v_n = 0 \tag{31}$$

by the inversion of the triangular matrix M_{n-1}.

Note that if we multiply (29) by M_n' on the left we have precisely the pseudo-normal equations (5) written in matrix form. Thus (29) and (5) are equivalent.

Now define the vectors

$$\mu_i = \begin{pmatrix} \mu_{i0} \\ \mu_{i1} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{i,\ i-1} \end{pmatrix} .$$

The equations for determining the a's and μ's can be written

$$v_{i+1} = a_{i+1} + A_i \ \mu_{i+1} . \qquad (32)$$

These give

$$A_i' \ v_{i+1} = A_i' \ A_i \ \mu_{i+1} ,$$

since all columns of A_i are orthogonal to a_{i+1}, or

$$A_i' \ v_{i+1} = D_i \ \mu_{i+1} ,$$

whence

$$\mu_{i+1} = D_i^{-1} \ A_i' \ v_{i+1} . \qquad (33)$$

In particular,

$$\mu_n = D_{n-1}^{-1} \ A_{n-1}' \ v_n \qquad (34)$$

so that equation (31) becomes

$$\sigma_n + M_{n-1}^{-1} \mu_n = 0. \qquad (35)$$

To summarize, we calculate μ_{i+1} and a_{i+1} sequentially by means of (32) and (33). We adjoin a_{i+1} to A_i to obtain A_{i+1}; we border M_i by μ_{i+1} and a unit row-vector to obtain M_{i+1}. Each vector a_{i+1} is that component of v_{i+1} that is orthogonal to the space of the previous v's. If the vectors v_i satisfied strictly a difference equation of order n, then v_n would be a linear combination of the preceding v's and a_n would vanish. We could thus continue the process until we found a vanishing a. Since the components of the v's are subject to errors of measurement, it is not to be expected that any vector a will vanish strictly, but we may expect that for some n, a_n will be negligibly small. Having found such an a_n, there remains only the simple calculation (35) to obtain σ_n.

There remains only one unsettled point; the choice of γ. Since all vectors v must have the same number of components, we must not choose γ so large that we run out of components as we include higher v's. Hence we must set some upper limit to the likely value of n, and choose γ accordingly. If γ is chosen too small we waste some of the experimental values at this stage. However, these can be picked up again when we improve our solution by the method of equations (16).

The above method is readily adapted, with minor modifications, to the fitting of general multiple-hit survival curves of the form

$$y = e^{-\alpha t}(1+\beta_1 t+\beta_2 t^2 + \dots + \beta_n t^n). \quad (36)$$

As N. M. Smith has pointed out, curves of this type are to be expected from irradiated organisms if there is a region within which n hits will be lethal, and also other regions in which a smaller number will suffice. Here t stands for dose.

This function satisfies an n-th order difference equation whose characteristic equation has all its roots equal. Hence the coefficients s_r determined as above must be such that for some u

$$s_r = (-)^r \binom{u}{r} u^r. \quad (37)$$

Consequently

$$\binom{n}{r} s_{r+1} = \binom{n}{r+1} s_r u \quad (38)$$

The consistency of these equations may provide a rough test of the hypothesis, but not a rigorous test since, of course, the measurements are not correctly weighted.

Having found u, one obtains α as before and then from (36) the β's can be found by ordinary least squares.

REFERENCES

1. E. Czuber. Wahrscheinlichkeitsrechnung. Leipzig and Berlin, 1908.
2. W. E. Deming. Statistical adjustment of data. New York, 1946.
3. Whittaker and Robinson. The Calculus of Observations. London and Glasgow, 1946.
4. F. A. Willers (translated by R. T. Beyer), Practical analysis. New York, 1948.