On the Best Fit to Experimental Data Curves Using Homogeneous Two-Parameter Functions

By R. W. Preisendorfer and B. W. Roos

1. Introduction. The following specific problem in fitting a theoretical curve to a given set of data forms the basis for the general method discussed in the present note: Some experimental data is given in the form of a discrete function g on an interval $[a, b]$ of the real axis. It is known that the data is associated with a phenomenon which in principle is represented analytically by the function f whose value at each point x of $[a, b]$ is given by $(\beta - \alpha x)/\ln (\beta/\alpha x)$. It is required to find the values of the two parameters α and β such that the theoretical curve is as close as possible, in the least square sense, to g.

The most direct approach to the solution of this problem is to choose a value of α and β and to compute the number

$$(\text{*}) \qquad G(\alpha, \beta) = \int_a^b \left[g(x) - \frac{\beta - \alpha x}{\ln (\beta/\alpha x)} \right]^2 dx.$$

This process is then repeated a number of times which results in a set $\{G(\alpha_1, \beta_1), \cdots, G(\alpha_n, \beta_n)\}$ of values of the integral (*). This set may then be inspected for a possible minimum. After some experimentation, a pair (α_0, β_0) may be found at which the function G attains a minimum. We then say that the function whose value at each x in $[a, b]$ is $(\beta_0 - \alpha_0 x)/\ln (\beta_0/\alpha_0 x)$ best approximates the given function g on $[a, b]$ in the least square sense.

Now the direct approach outlined above can be considerably shortened and made somewhat less hazardous ("Have I really found the pair (α, β) which minimizes G?") by observing the following interesting property of the given theoretical function f: For each x in $[a, b]$

$$\frac{\beta - \alpha x}{\ln \left(\dfrac{\beta}{\alpha x} \right)} = \beta \left(\frac{1 - \dfrac{\alpha x}{\beta}}{\ln \left(\dfrac{\beta}{\alpha x} \right)} \right) = \beta \left(\frac{tx - 1}{\ln tx} \right)$$

where we have set $t = \alpha/\beta$. The immediate consequence of this observation is that we have factored one of the parameters out of the theoretical function, leaving a new function which is the product of β (one of the original parameters) and a function which has only one parameter (namely t). One would suspect that the trial and error procedure of the direct approach could be appreciably shortened by capitalizing on this fact. We shall now show that this is indeed the case. The net result is a general method of finding the best fitting two-parameter function by means of one-parameter techniques, i.e., using techniques that allow a determination of the required pair (α_0, β_0) *one member at a time* instead of two at a time as in the preceding direct approach. Before going into the details of the method, it is of interest to see how large a domain of two-parameter functions is accessible to this method.

2. Semi-Homogeneous Two-Parameter Functions. The theoretical function f discussed above is a special case of a two-parameter function $f(\,\cdot\,;\,\alpha,\,\beta)$ with the property that for each x in $[a,\,b]$,

(1)   I. $f(x;\alpha,\beta) = \alpha^n F(x;t), \qquad t = \beta/\alpha, \qquad n \geqq 0$

or with the property:

(2)   II. $f(x;\alpha,\beta) = \beta^n F(x;t), \quad t = \alpha/\beta, \quad n \geqq 0$

where the domain of parameters of F is the real line. A function $f(\,\cdot\,;\,\alpha,\,\beta)$ on an interval $[a,\,b]$ with property I or II is said to be a *semi-homogeneous function of two parameters* (of degree n). Two functions are *similarly homogeneous* if they are both of type I or both of type II.

Some examples of semi-homogeneous functions of two parameters are:

$$f(x;\alpha,\beta) = (\beta - \alpha x)^n/\ln\,(\beta/\alpha x), \qquad\qquad n \geqq 0$$

$$f(x;\alpha,\beta) = (\alpha + \beta x)^n, \qquad\qquad n \geqq 0$$

$$f(x;\alpha,\beta) = (\alpha + \beta x)^n \sin\,(\alpha x/\beta), \qquad\qquad n \geqq 0$$

$$f(x;\alpha,\beta) = (\alpha + \beta x)^n \exp\,(\alpha x/\beta), \qquad\qquad n \geqq 0$$

$$f(x;\alpha,\beta) = (\alpha + \beta x)^n F(x;t), \qquad t = \alpha/\beta, \qquad n \geqq 0$$

New examples of semi-homogeneous functions can be manufactured from old ones by observing that under suitable conditions the sum, difference, and product of two similarly homogeneous two-parameter functions (with the same parameters) is again a semi-homogeneous function of the same type. Thus, e.g., if $f(\,\cdot\,;\,\alpha,\,\beta)$ and $g(\,\cdot\,;\,\alpha,\,\beta)$ are similarly homogeneous of type I and of degree m and n,

$$h(x;\alpha,\beta) = f(x;\alpha,\beta)g(x;\alpha,\beta)$$

$$= \alpha^m F(x;t)\alpha^n G(x;t)$$

$$= \alpha^{m+n} H(x;t)$$

where

$$H(x;t) = F(x;t)G(x;t).$$

If f and g are of the same degree their sum (or difference) is again a semi-homogeneous two-parameter function of the same type and degree. These algebraic properties are summarized below: Every set of similarly homogeneous functions of the same parameters is closed under the operations of

a) Sum
b) Difference
c) Multiplication
d) Division (when possible)
e) Multiplication by a scalar when the functions are of the same degree;

A set is closed under the operations of:

a) Multiplication
b) Division (when possible)
c) Multiplication by a scalar—when the functions are of *arbitrary degree*

Finally, the sum of two semi-homogeneous functions of dissimilar type but of *same degree*, is semi-homogeneous of one or the other type.

One is now naturally led to ask: what happens if two semi-homogeneous functions of dissimilar type are multiplied together? In this way we come to the concepts of the following section.

3. Homogeneous Two-Parameter Functions. To answer the question raised in the preceding section, consider $f(\,\cdot\,;\alpha,\beta)$ and $g(\,\cdot\,;\alpha,\beta)$ which are of type I and II respectively. Then:

$$
\begin{aligned}
h(x;\alpha,\beta) &= f(x;\alpha,\beta)g(x;\alpha,\beta) \\
&= \alpha^n F\left(x;\frac{1}{t}\right)\beta^m G(x;t), \qquad t = \alpha/\beta \\
&= \alpha^n \beta^m H(x;t)
\end{aligned}
$$

where

$$
H(x;t) = F\left(x;\frac{1}{t}\right) G(x;t).
$$

We will then say that a two-parameter function $f(\,\cdot\,;\alpha,\beta)$ on $[a,b]$ is *homogeneous of degree m in α and degree n in β* if for each x in $[a,b]$,

$$
(3) \qquad f(x;\alpha,\beta) = \alpha^m \beta^n F(x;t), \qquad t = \alpha/\beta \qquad m,n \geqq 0.
$$

From this definition, the following properties of homogeneous two-parameter functions are easily deduced:

The product of two homogeneous two-parameter functions of arbitrary degrees in the parameters is again homogeneous; if they are of the same degree in each parameter their sum and difference is homogeneous and of the same degree in each parameter. In any event, multiplication by a scalar preserves homogeneity.

Perhaps the most important observations at this point of the discussion are first that the set of all *semi*-homogeneous two-parameter functions on $[a,b]$ is a proper subset of the set of all *homogeneous* two-parameter functions on $[a,b]$; and secondly, that the main problem of the best fit to a given curve can be solved with comparable ease in the more general homogeneous context as in the semi-homogeneous context. We now turn to the details of the method.

4. Derivation Of The Method. Let g be a given function on an interval $[a,b]$. Here g represents, for example, a set of given physical data correlated with the numbers of $[a,b]$. Then g is to be approximated by a member of some given family of homogeneous two-parameter functions $f(\,\cdot\,;\alpha,\beta)$ on $[a,b]$ (of fixed degrees $m \geqq 0$ in α and $n \geqq 0$ in β such that not both m and n are zero) such that the function G defined for each pair (α,β) of parameters by:

$$
\begin{aligned}
G(\alpha,\beta) &= \int_a^b [g(x) - f(x;\alpha,\beta)]^2\, dx \\
&= \int_a^b [g(x) - \alpha^m \beta^n F(x;t)]^2\, dx, \qquad t = \alpha/\beta
\end{aligned}
$$

attains a minimum for some pair (α_0, β_0). It will be assumed that the required parameters α_0 and β_0 are not zero so that the calculations below can be carried out without ambiguity. The singular possibilities ($\alpha_0 = 0$ or $\beta_0 = 0$) can be examined to see if they yield a solution *after* the pair has been found by the procedure now to be described.

A necessary condition that G attains a minimum at (α_0, β_0) is that

$$\frac{\partial G(\alpha_0, \beta_0)}{\partial \alpha} = 0 \qquad \frac{\partial G(\alpha_0, \beta_0)}{\partial \beta} = 0.$$

In more detail, these conditions require:

$$\int_a^b [g(x) - \alpha_0^m \beta_0^n F(x; t_0)] \left[-m\alpha_0^{m-1}\beta_0^n F(x; t_0) - \alpha_0^m \beta_0^n \frac{\partial F(x; t_0)}{\partial t} \cdot \frac{1}{\beta_0} \right] dx = 0$$

$$\int_a^b [g(x) - \alpha_0^m \beta_0^n F(x; t_0)] \left[-n\alpha_0^m \beta_0^{n-1} F(x; t_0) + \alpha_0^m \beta_0^n \frac{\partial F(x; t_0)}{\partial t} \cdot \frac{\alpha_0}{\beta_0^2} \right] dx = 0.$$

Multiplying the first equation by α_0/β_0, and adding the result to the second, it follows that

$$(m + n)\alpha_0^m \beta_0^{n-1} \int_a^b [g(x) - \alpha_0^m \beta_0^n F(x; t_0)]F(x; t_0)\, dx = 0.$$

By the standing agreement on the values of m, n, α_0, β_0, it follows that the integral must be zero, so that

$$(4) \qquad \alpha_0^m \beta_0^n = \frac{\displaystyle\int_a^b g(x)F(x; t_0)\, dx}{\displaystyle\int_a^b F^2(x; t_0)\, dx} .$$

This is the key theoretical relation for the present method. To see the role it plays, we assume, as is suggested by (4), that the product $\gamma = \alpha^m \beta^n$ is a function of t. Then the general expression for $G(\alpha, \beta)$ can be written as

$$H(t) = \int_a^b [g(x) - \gamma(t)F(x; t)]^2\, dx$$

which unfolds into:

$$H(t) = \int_a^b g^2(x)\, dx - 2\gamma(t) \int_a^b g(x)F(x; t)\, dx + \gamma^2(t) \int_a^b F^2(x; t)\, dx$$

making use of the definition

$$(5) \qquad \gamma(t) = \frac{\displaystyle\int_a^b g(x)F(x; t)\, dx}{\displaystyle\int_a^b F^2(x; t)\, dx}$$

we deduce that

$$(6) \qquad H(t) = \int_a^b g^2(x)\, dx - \gamma^2(t) \int_a^b F^2(x; t)\, dx \geqq 0.$$

It follows that $H(t)$ is a minimum when the quantity

$$J(t) \equiv \gamma^2(t) \int_a^b F^2(x; t)\, dx \geqq 0$$

is a maximum, i.e., when

(7)
$$J(t) = \frac{\left[\int_a^b g(x)F(x, t)\, dx\right]^2}{\int_a^b F^2(x; t)\, dx}$$

is a maximum.

The problem of minimizing $G(\alpha, \beta)$ by varying the *two* parameters α and β simultaneously has now been reduced to the problem of maximizing $J(t)$ by varying the *one* parameter t. For when the t_0 which maximizes (7) is found, we return to (4) and compute $\alpha_0{}^m \beta_0{}^n = \gamma_0$ as shown. By noting that $\gamma_0 = t_0{}^m \beta_0{}^{m+n}$, we can then find $\beta_0{}^{m+n}$ and all the (possibly complex) $(m + n)$th roots of $\beta_0{}^{m+n}$. Then for each root β_0 we compute the corresponding α_0 by means of the relation $\alpha_0 = t_0 \beta_0$. The physical or mathematical context of the particular curve-fittin9 problem at hand should then lead to the choice of the appropriate pairs (α_0, β_0).

5. Summary Of The Method.

Given: An empirical function g on an interval $[a, b]$, and a family of homogeneous two-parameter functions $f(\,\cdot\,; \alpha, \beta)$ on $[a, b]$ of degree $m \geqq 0$ in α and $n \geqq 0$ in β such that not both m and n are zero.

Required: The values α_0 and β_0 of α and β such that

$$G(\alpha, \beta) = \int_a^b [g(x) - f(x; \alpha, \beta)]^2\, dx \geqq G(\alpha_0, \beta_0)$$

for all (α, β).

The required pair (α_0, β_0) is found as follows:

1. Compute $J(t)$ over some range where

$$J(t) = \frac{\left[\int_a^b g(x)F(x; t)\, dx\right]^2}{\int_a^b F^2(x; t)\, dx}$$

and where $F(x; t)$ is defined by:

$$f(x; \alpha, \beta) = \alpha^m \beta^n F(x; t)$$

2. Find t_0 which maximizes $J(t)$. Let this maximum be $J(t_0)$.
3. Compute

$$\gamma_0 = J(t_0) \Big/ \left[\int_a^b g(x)F(x; t_0)\, dx\right].$$

4. Compute

$$\beta_0{}^{m+n} = \gamma_0 / t_0{}^m$$

and then β_0 (in either real or complex form as required by the problem at hand).

5. Compute

$$\alpha_0 = t_0\beta_0$$

6. Compare $G(\alpha_0, 0)$, $G(0, \beta_0)$, $G(0, 0)$, and $G(\alpha_0, \beta_0)$ and choose that pair: $(\alpha_0, 0)$, $(0, \beta_0)$, $(0, 0)$, (α_0, β_0) which minimizes $G(\alpha, \beta)$.

Observations: The last step 6 is needed to complete the procedure since the possibilities $\alpha_0 = 0$ and $\beta_0 = 0$ where excluded from the analysis leading to steps 1–5. Finally, it may be that a particular $F(x; t)$ associated with a homogeneous two-parameter function is such that step 2 yields more than one value of the parameter t_0. In this case the method yields several additional pairs (α_0, β_0) which must be examined for the minimizing property.

John Jay Hopkins Laboratory for Pure and Applied Science
General Atomic Division of General Dynamics Corporation
San Diego, California