A general approach to simultaneous curve fitting

by David Smith

National Physical Laboratory, Teddington, United Kingdom

1. Introduction

The report BIPM-79/12 by J W Müller[1] discussed the method of simultaneous curve fitting, in which several polynomials are fitted simultaneously with the condition that all the polynomials have a common value when the dependent variable has the value zero. The purpose of this note is to point out that this process can be regarded as a special case of fitting a single multidimensional surface through all the data points. The great advantage is that all the unknown parameters and their variances can then be readily obtained with a standard linear least squares computer routine which would be available at most computer installations.

2. The multidimensional approach

Let there be n polynomial functions represented by

$$y = a_1 + f_k(x) , \qquad \text{for } k = 1,2 \ldots n ,$$

where $f(x) \rightarrow o$ as $x \rightarrow o$

and where a_1 is the common intercept with the y axis.

These n functions can be considered as a single multidimensional surface given by

$$y = a_1 + \sum_{k=1}^{n} f_k(x_k) \equiv a_1 + g(x_1, x_2, \ldots x_n),$$

whose intercept with the y axis is the desired common intercept. All the data points are confined to the various $(y - x_k)$ planes, so the data for the kth line $y = a_1 + f_k(x)$ is expressed as $y = a_1 + g(0, 0 ,,, x_k, ,,,0)$ where all the x parameters except the kth, are zero.

Since a polynomial function is linear in its coefficients, the multidimensional surface is also a function linear in its coefficients, and can be written as

$$y = \sum_j a_j Z_j \quad ,$$

where the a_j represent the coefficients (i.e. parameters to be fitted) and where the Z_j represent the known x data.

For example, two quadratic functions with a common intercept is expressed as

$$y = a_1 + a_2x_1 + a_3x_1^2 + a_4x_2 + a_5x_2^2 = \sum_{j=1}^{5} a_j Z_j \quad .$$

Thus

$$Z_1 = 1.0,\ Z_2 = x_1,\ Z_3 = x_1^2,\ Z_4 = x_2,\ Z_5 = x_2^2 \ .$$

In general, the sum of squares is

$$\sum_i w_i(y_i - \sum_j a_j Z_{ji})^2 \quad ,$$

where the outer sum is over all the data points (w_i, y_i, Z_{ji}).

Regarded in this manner, all the unknown parameters (a_j) and their variances and covariances can be readily obtained with a standard least squares computer routine (i.e. for multiple linear regression of a dependent variable y on n dependent variables $Z_1 \ldots Z_n$) which would be available as a "library program" at most computer installations.

Example: Three lines with common intercept.

This example is used in BIPM-79/12 and is illustrated in figure 1, where the 3 lines are given by

1st line: $y = a_1 + a_2x + a_3x^2 + a_4x^3$

2nd line: $y = a_1 + a_5x + a_6x^2 + a_7x^3$

3rd line: $y = a_1 + a_8x + a_9x^2 + a_{10}x^3$

and where a_1 is the common intercept.

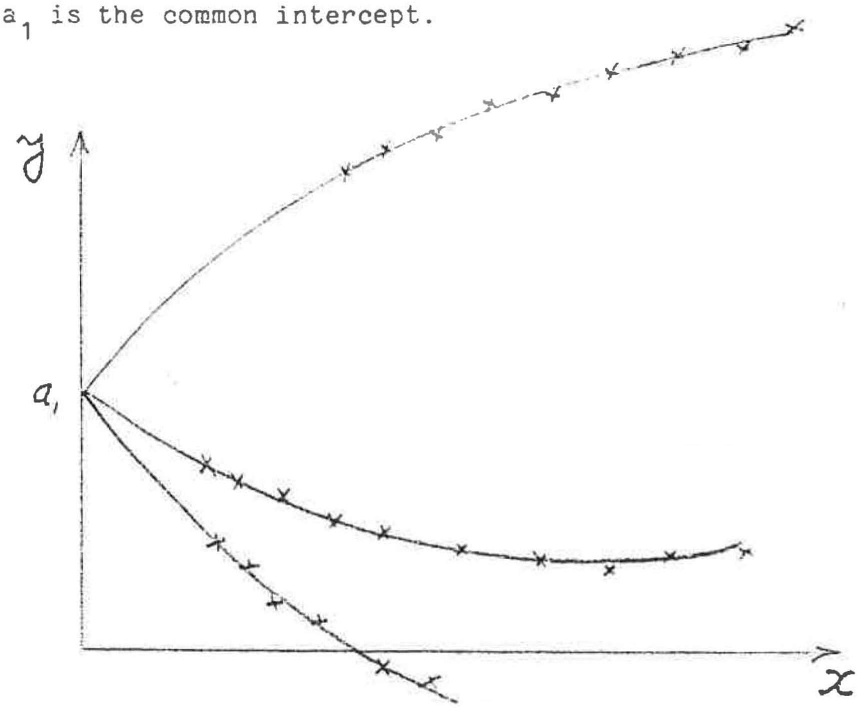

Figure 1: Three cubic lines with common intercept.

In the multidimensional representation, shown in figure 2,

$$\begin{aligned} y = a_1\, 1.0 \quad & + a_2x_1 + a_3x_1^2 + a_4x_1^3 \\ & + a_5x_2 + a_6x_2^2 + a_7x_2^2 \\ & + a_8x_3 + a_9x_3^2 + a_{10}x_3^2 \end{aligned}$$

$$\equiv \sum_{j=1}^{10} a_j Z_j \ .$$

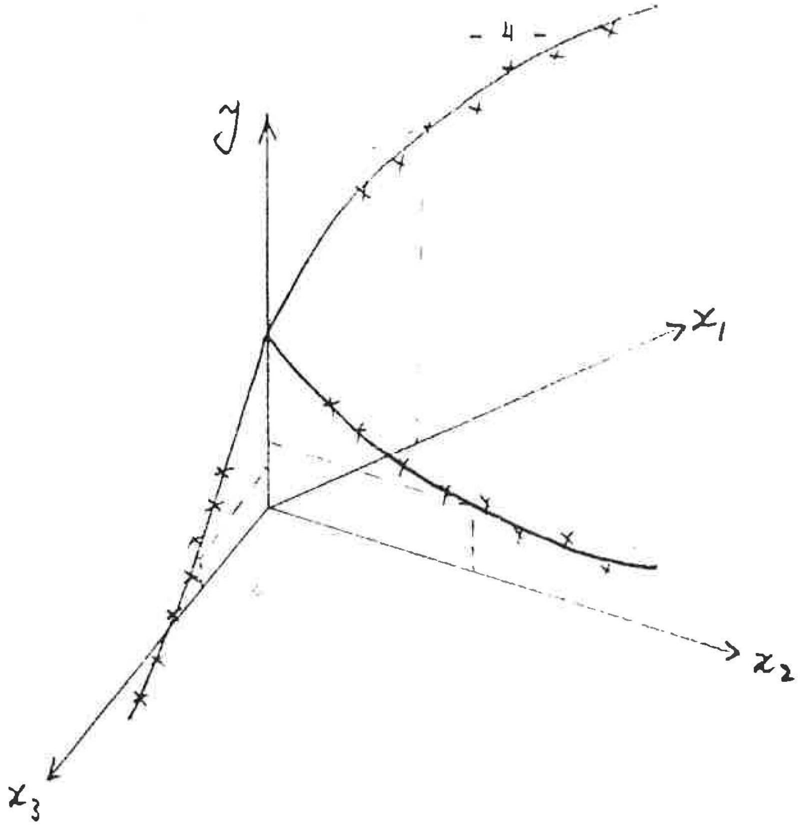

Figure 2: Three cubic lines with common intercept.

The data for the 3 lines all lie along the various ($y-x_k$) planes, and can be expressed as follows

	y	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	Z_{10}
1st line	y	1.0	x	x^2	x^3	0	0	0	0	0	0
2nd line	y	1.0	0	0	0	x	x^2	x^3	0	0	0
3rd line	y	1.0	0	0	0	0	0	0	x	x^2	x^3

The sum of squares is $\sum_i w_i(y_i - \sum_j a_j Z_{ji})^2$ where the sum is taken over all the separate data points for all three lines. This data is now in the standard format and can be fitted to the multidimensional surface by the standard least squares routine, in this case with 10 dependent variables (Z_1 ... Z_{10}) to give the 10 parameters (a_1 ... a_{10}).

3. Extension to fit several functions of 2 independent x variables

Suppose there are n functions to be fitted

$$y = a_1 + f_k(xa,xb) \quad , \qquad \text{for } k = 1,2 \ldots n \; ,$$

where $f(xa,xb) \rightarrow o$ as $xa,xb \rightarrow o$

and where a_1 is the common intercept with the y axis.

These functions can be regarded as a single multidimensional function given by

$$y = a_1 + \sum_{k=1}^{n} f_k(xa_k,\ xb_k) \equiv a_1 + g(xa_1,xb_1,,,xa_n,xb_n)$$
$$\equiv \sum_j a_j Z_j \; .$$

Again the data points are confined entirely to the various (y,xa_k,xb_k) spaces.

Example: Two surfaces with common intercept.

Suppose each surface is a function of 2 independent variables xa (linear) and xb (quadratic) :

1st surface: $y = a_1 + a_2xa + a_3xb + a_4xb^2$,

2nd surface: $y = a_1 + a_5xa + a_6xb + a_7xb^2$.

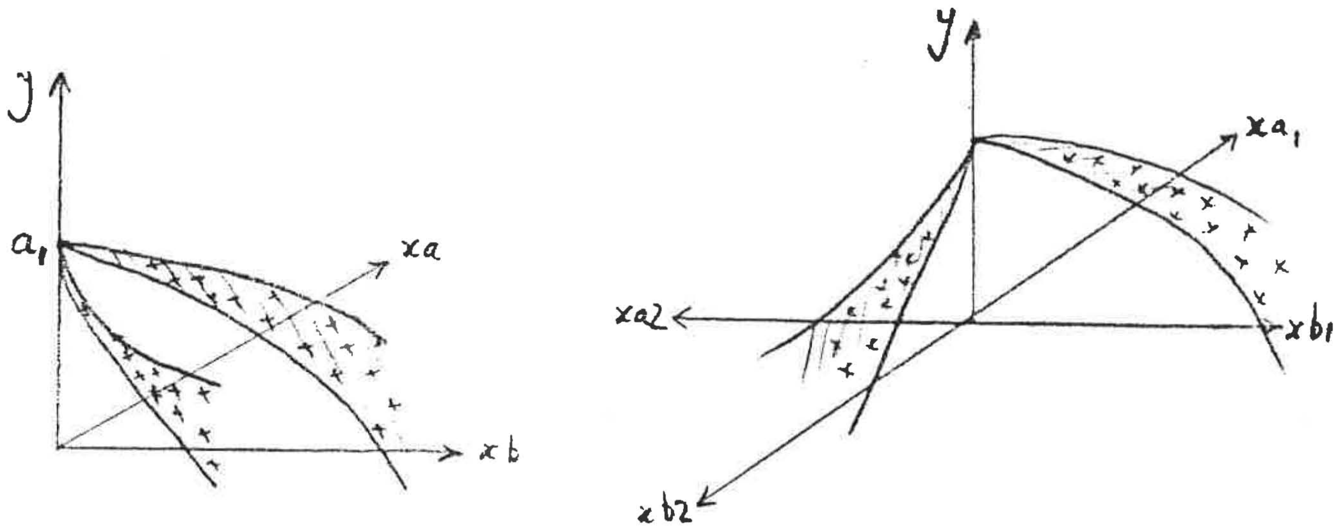

Figure 3: Two surfaces with common intercept.

In the multidimensional representation (figure 3)

$$y = a_1\, 1.0 + a_2 x_1 + a_3 x_2 + a_4 x_2^2 + a_5 x_3 + a_6 x_4 + a_7 x_4^2 = \sum_{j=1}^{7} a_j Z_j \ .$$

The data can be expressed as

	y	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7
1st surface	y	1.0	xa	xb	xb^2	0	0	0
2nd surface	y	1.0	0	0	0	xa	xb	xb^2

This data is now in the standard format and can be fitted by the same least squares program, in this case with 7 unknowns $a_1 .. a_7$.

4. Extension to include other constraints

Although not directly relevant to problems of $4\pi\beta - \gamma$ extrapolation, this method can be used to fit various functions simultaneously with various constraint(s) such as common gradients and/or intercepts.

Example: 3 quadratic lines, one pair with common intercept, one pair common gradient parameters.

1st line: $y = a_1 + a_2 x + a_3 x^2$

2nd line: $y = a_1 + a_4 x + a_5 x^2$

3rd line: $y = a_6 + a_4 x + a_5 x^2$

(1st and 2nd lines:) common intercept a_1

(2nd and 3rd lines:) common gradients a_4, a_5

In the multidimensional representation,

$$y = a_1x_1 + a_2x_2 + a_3x_2^2 + a_4x_3 + a_5x_3^2 + a_6x_4 = \sum_{j=1}^{6} a_jZ_j .$$

The two intercepts are represented by $a_1x_1 + a_6x_4$, where for lines 1 and 2, $x_1 = 1$, $x_4 = 0$, (intercept a_1) whereas for line 3, $x_1 = 0$, $x_4 = 1$ (intercept a_6). The common gradient parameters are given by a_4 and a_5. The data is expressed as follows

	y	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
1st line	y	1.0	x	x^2	0	0	0
2nd line	y	1.0	0	0	x	x^2	0
3rd line	y	0	0	0	x	x^2	1.0

The data is once again in the standard format and can be fitted by the <u>same</u> least squares program, in this case with 6 unknowns $a_1, \ldots, a_6$.

Conclusion

The technique described enables a range of fitting problems to be solved with one and the same standard computer routine (i.e. multiple linear regression of a dependent variable y on n independent variables). In principle any number of functions of any type can be used, provided such functions are linear in the unknown parameters.

An extension to the use of functions non-linear in the parameters is described in the appendix.

References 1) J W Müller: "Simultaneous curve fitting", Rapport BIPM-79/12.

Appendix

Extension to functions non-linear in the parameters to be fitted.

The ideas described in the paper can be applied to more complicated functions which are non-linear in the unknown parameters, provided a suitable general non-linear least squares minimisation computer routine is available.

For example, consider the functions

$$y = a_1(1 + bx + cx^2) ,$$

$$y = a_2(1 + bx + cx^2) ,$$

$$y = a_3(1 + bx + cx^2) .$$

These represent 3 lines with different intercepts with the y axis, but with common parameters b,c.

In a multidimensional representation, this is equivalent to fitting the following function,

$$y = (a_1x_1 + a_2x_2 + a_3x_3)(1 + a_4x_4 + a_5x_4^2) ,$$

where y is a non-linear function of the 5 unknown parameters $a_1 \ldots a_5$, and where the intercepts are given by $a_1x_1 + a_2x_2 + a_3x_3$. The data is expressed as

	y	x_1	x_2	x_3	x_4
1st line	y	1.0	0	0	x
2nd line	y	0	1.0	0	x
3rd line	y	0	0	1.0	x

The sum of squares is

$$\sum w_i \left[y_i - (a_1x_1 + a_2x_2 + a_3x_3)(1 + a_4x_4 + a_5x_4^2) \right]^2 .$$

Clearly, this requires a computer routine for non-linear minimisation with respect to the parameters $a_1, \ldots, a_5$.

(August 1981)