The Estimation of Parameters in Regression Functions Subject to Certain Restraints

PAUL D. MINTON AND ALFRED E. CROFTS, JR.[1]

1. *Introduction.* We consider two types of problems in maximum likelihood estimation of parameters of linear functions subject to certain restraints. One is a family of lines with equal slopes or intercepts; the other is a pair of lines constrained to meet at a predetermined point.

In the case of normally distributed errors with equal variances within each set, the solutions are identical with least squares solutions. In addition to linear functions, non-linear functions which are transformable to linearity may be treated under these methods.

2. *Background.* Quandt [2][2] estimated the point of separation of two lines in a regression system. In the case of a known point of separation, he gave methods for estimating parameters, but did not require that the lines meet at the given point. Tocher [3] gave a test for equal intercepts on several lines. Page [1] gave a test for a change in a parameter occurring at an unknown point and Quandt also proposed a test that no switch occurred. Williams [4, 5] discussed some general problems relating to estimation in simultaneous equations.

3. *Notation.* Given k functions of the form

$$y = \alpha_i + \beta_i x, \qquad i = 1, 2, \ldots, k, \tag{1}$$

(or transformable to this form) and a set of observations (x_{ij}, y_{ij}), $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n_i$; where (x_{ij}, y_{ij}) is the j-th observation in the i-th set, the problem is to estimate the α_i, β_i, and the σ_i^2, subject to certain restraints.

Suppose the data are related by

$$y_{ij} = \alpha_i + \beta_i x_{ij} + u_{ij}, \tag{2}$$

where u_{ij} is normally distributed with mean zero and variance σ_i^2. We assume the x's measured without error. Then

$$f(u_{ij}) = f(y_{ij} - \alpha_i - \beta_i x_{ij}) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2\sigma_i^2} u_{ij}^2\right]. \tag{3}$$

The likelihood functions are

$$L_i = \prod_{j=1}^{n_i} f(u_{ij}) \qquad \text{and } L = \prod_{i=1}^{k} L_i. \tag{4}$$

To estimate the parameters by maximum likelihood, set

$$\frac{\partial L}{\partial \alpha_i} = 0, \frac{\partial L}{\partial \beta_i} = 0, \text{ and } \frac{\partial L}{\partial \sigma_i^2} = 0, \qquad i = 1, 2, \ldots, k. \tag{5}$$

---

[1]Department of Mathematics, Southern Methodist University.

[2]Square brackets refer to the bibliography.

We impose some restraint on the α's and β's denoted generically by $g(\alpha, \beta, x, y) = 0$, either by adjusting (2) and hence (5), or by differentiating $L - \lambda g$ with respect to the appropriate α_i, β_i, σ_i^2 and λ, where λ is a Lagrange multiplier.

4. *A family of lines with equal slopes or intercepts.* In this section, summation over j is from 1 to n_i and summation over i is from 1 to k.

For the case of equal slopes, $\beta_1 = \beta_2 = \ldots = \beta_k = \beta$, and (4) becomes

$$L_i = \left[\frac{1}{\sqrt{2\pi\,\sigma_i^2}}\right]^{n_i} exp\left[-\frac{1}{2\,\sigma_i^2}\sum_j (y_{ij} - \alpha_i - \beta x_{ij})^2\right]. \qquad (6)$$

Then (5) gives

$$\hat{\beta} = \frac{\sum_i \frac{1}{\sigma_i^2}\left[\sum_j x_{ij}\,y_{ij} - \frac{(\sum_j x_{ij})\,(\sum_j y_{ij})}{n_i}\right]}{\sum_i \frac{1}{\sigma_i^2}\left[\sum_j x_{ij}^2 - \frac{(\sum_j x_{ij})^2}{n_i}\right]}, \qquad (7)$$

$$\hat{\alpha}_i = \frac{\sum_j y_{ij} - \hat{\beta}\sum_j x_{ij}}{n_i} = \bar{y}_i - \hat{\beta}\,\bar{x}_i, \qquad (8)$$

and $$\hat{\sigma}_i^2 = \frac{1}{n_i}\sum_j (y_{ij} - \hat{\alpha}_i - \hat{\beta}\,x_{ij})^2,\ i = 1, 2, \ldots, k. \qquad (9)$$

Note that when the σ_i^2 are equal, the results for α and β reduce to the least squares estimates, in which the σ's cancel, and the estimate of σ^2 becomes $\frac{1}{n}\sum_i\sum_j (y_{ij} - \hat{\alpha}_i - \hat{\beta}\,x_{ij})^2$, where $n = \sum_i n_i$.

For the case of equal intercepts, $\alpha_1 = \alpha_2 = \ldots = \alpha_k = \alpha$, *and* (4) becomes

$$L_i = \left(\frac{1}{\sqrt{2\pi\,\sigma_i^2}}\right)^{n_i} exp\left[-\frac{1}{2\,\sigma_i^2}\sum_j (y_{ij} - \alpha - \beta_i\,x_{ij})^2\right].$$

Then (5) gives

$$\hat{\alpha} = \frac{\sum_i \frac{1}{\sigma_i^2}\left[\sum_j y_{ij} - \frac{(\sum_j x_{ij})\,(\sum_j x_{ij}\,y_{ij})}{\sum_j x_{ij}^2}\right]}{\sum_i \frac{1}{\sigma_i^2}\left[n_i - \frac{(\sum_j x_{ij})^2}{\sum_j x_{ij}^2}\right]},$$

$$\hat{\beta}_i = \frac{\sum_j x_{ij}\,y_{ij} - \hat{\alpha}\sum_j x_{ij}}{\sum_j x_{ij}^2},$$

and $\hat{\sigma}_i^2 = \frac{1}{n_i} \sum_j (y_{ij} - \hat{\alpha} - \hat{\beta}_i x_{ij})^2, \quad i = 1, 2, \ldots, k.$

5. *Two lines required to meet at $x = x_o$.* In this section, summations are all on j; for the subscript ij, summation is from $j = 1$ to n_i, for $i = 1, 2$.

We consider two lines, knowing which points belong to each set, and want them to meet at $x = x_o$. Then

$$L = \frac{1}{2\pi \sigma_1 \sigma_2} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma_1^2} \Sigma (y_{1j} - \alpha_1 - \beta_1 x_{1j})^2 + \frac{1}{\sigma_2^2} \Sigma (y_{2j} - \alpha_2 - \beta_2 x_{2j})^2\right]\right\}$$

and the constraint is $g(\alpha, \beta, x, y) = \alpha_1 + \beta_1 x_o - \alpha_2 - \beta_2 x_o = 0.$

Differentiating $L - \lambda g$ with respect to $\alpha_1, \beta_1, \alpha_2, \beta_2$ and λ and eliminating λ gives a set of normal equations to solve for these parameters:

$$(\sigma_2^2 n_1)\,\alpha_1 + (\sigma_1^2 n_2)\,\alpha_2 + (\sigma_2^2 \Sigma x_{1j})\,\beta_1 + (\sigma_1^2 \Sigma x_{2j})\,\beta_2 = \sigma_2^2 \Sigma y_{1j} + \sigma_1^2 \Sigma y_{2j}$$

$$(\Sigma x_{1j} + x_o n_1)\,\alpha_1 + (\Sigma x_{1j}^2 + x_o \Sigma x_{1j})\,\beta_1 = \Sigma x_{1j} y_{1j} - x_o \Sigma y_{1j}$$

$$(\sigma_2^2 x_o n_1)\,\alpha_1 + (\sigma_1^2 \Sigma x_{2j})\,\alpha_2 + (\sigma_2^2 x_o \Sigma x_{1j})\,\beta_1 + (\sigma_1^2 \Sigma x_{2j}^2)\,\beta_2 = \sigma_1^2 \Sigma x_{2j} y_{2j} + \sigma_2^2 x_o \Sigma y_{1j}$$

$$\alpha_1 - \alpha_2 + x_o \beta_1 - x_o \beta_2 = 0,$$

$$\hat{\sigma}_1^2 = \frac{1}{n_1} \Sigma (y_{1j} - \hat{\alpha}_1 - \hat{\beta}_1 x_{1j})^2,$$

$$\hat{\sigma}_2^2 = \frac{1}{n_2} \Sigma (y_{2j} - \hat{\alpha}_2 - \hat{\beta}_2 x_{2j})^2.$$

As before, if the σ's are equal, they cancel in the normal equations for the α's and β's and the estimate of σ^2 is $\frac{1}{n} \sum_{i=1}^{2} \sum_j (y_{ij} - \hat{\alpha}_i - \hat{\beta}_i x_{ij})^2.$

6. *Solutions to the normal equations.* If the σ_i^2 are not assumed to be equal, the system of normal equations which must be solved for the parameters in the above cases becomes non-linear. Then it may be possible to apply an iterative method of solution, perhaps with the aid of a digital computer.

As an example where this method was successful, we give the following data on vacuum tube characteristic curves, in this case three lines assumed parallel. Here, x is D.C. load (MA), and y is D.C. volts to filter input. The three lines corresponded to three voltages.

Vacuum Tube Characteristics

Line	*Volts*	x	y
1	220	40	170
		80	150
		100	145
		120	140
2	250	40	195
		80	175
		100	170
		120	165
3	400	40	325
		80	315
		100	305
		120	295

To start the solution, β was estimated graphically and substituted into (8) with the data on x_i and y_i to produce estimates of the α_i; the $\hat{\alpha}_i$ and $\hat{\beta}$ were then used in (9) to give a set of $\hat{\sigma}_i^2$, which were then used in (7) to produce a new estimate of β. The process was then repeated. The estimates converged as shown in the following table:

		$\hat{\beta}$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$
Graphical estimate:		—.3	180	205	335	20	28	9
Iterations:	1	—.351	180	205.1	338.8	6.24	8.32	6.62
	2	—.391	184.5	209.5	343.2	4.42	2.82	4.62
	3	—.402	185.4	210.4	344.2			

7. *Some applications.* Examples of a family of lines with equal slopes or intercepts are budget lines in economics (straight lines), and vacuum tube characteristic curves in electrical engineering (a family of exponentials with equal exponential parameters for which the logarithms are linear).

The case of two lines constrained to meet at a given point has occurred in economics as gasoline consumption curves, breaking at a known date, and as reaction curves in a chemical experiment, breaking at a "bubble point."

Bibliography

[1] PAGE, E. S. A test for a change in a parameter occurring at an unknown point, Biometrika, 42, 1955, 523-7.
[2] QUANDT, RICHARD E. The estimation of the parameters of a linear regression system obeying two separate regimes. Journal of the American Statistical Association, 53. 1958, 873-80.
[3] TOCHER, K. D. On the concurrence of a set of regression lines. Biometrika, 39, 1952, 109-117.
[4] WILLIAMS, E. J. *Regression Analysis.* John Wiley and Sons, New York, 1959.
[5] ———. Tests of significance for concurrent regression lines. Biometrika, 40, 297-305.