Recall that SL(2, 7) is the special linear group of 2×2 matrices of determinant 1 over the integers modulo 7. Also, PSL(2, 7) is the factor group of SL(2, 7) by $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right\}$, and PSL(2, 7) has order 168.

EXERCISE. Use a counting argument to prove PSL(2, 7) is simple.

References

1. John R. Durbin, Modern Algebra: An Introduction, Wiley, New York, 1979.
2. John B. Fraleigh, A First Course in Abstract Algebra, 2nd ed. Addison-Wesley, Reading, MA, 1976.
3. William J. Gilbert, Modern Algebra with Applications, Wiley, New York, 1976.
4. I. N. Herstein, Topics in Algebra, Xerox, Lexington, MA, 1975.
5. John Rose, A Course on Group Theory, Cambridge University Press, Cambridge, 1978.
6. Dan Saracino, Abstract Algebra: A First Course, Addison-Wesley, Reading, MA, 1980.
7. W. R. Scott, Group Theory, Prentice-Hall, Englewood Cliffs, NJ, 1964.

A GENERALIZED PROBLEM OF LEAST SQUARES

HUNG C. LI

Department of Mathematics, University of Southern Colorado, Pueblo, CO 81001

In regression analysis [1], the process of determining an interpolation function for a given set of data under the least squares criterion, one assumes the independent variables to be exact and not subject to error or variation. For instance, in elementary statistics one usually uses vertical deviations. In this context x is called the independent variable which is under control and assumed to be exact (occasionally, some textbooks [2] mention using horizontal deviations; in this situation x and y interchange roles). If one allows all the variables to reflect error, the problem of determining the "best-fitting" function for a given set of data becomes more complicated. In this paper I will consider this generalized least squares problem for the determination of an interpolation function in two variables in detail.

Consider a set of n points $(x_1, y_1), \ldots, (x_n, y_n)$. Without loss of generality, one can assume that the n points are noncollinear. The generalized linear interpolation function for the data is the line

(1) $$y - b - mx = 0,$$

such that the sum of squares of the perpendicular distances Σd_j^2 (j runs from 1 to n) from the given n points to the line (1) is a minimum, where

$$d_j = |y_j - b - mx_j| \Big/ \sqrt{1 + m^2}.$$

The object is to determine the parameters b and m in (1) which minimize Σd_j^2.

First note that x_j and y_j are the fixed numbers which one has observed. Denote

$$f(b, m) = \Sigma d_j^2 = \frac{1}{1 + m^2} \Sigma (y_j - b - mx_j)^2.$$

(2) $$\partial f / \partial b = 0 \text{ and } \partial f / \partial m = 0 \text{ give}$$
$$b = \bar{y} - m\bar{x},$$

and

(3) $$Am^2 - Bm - A = 0,$$

where $\bar{x}$ and $\bar{y}$ are the arithmetic means of $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ respectively, and

$$A = \Sigma x_j y_j - n\bar{x}\bar{y} \text{ and } B = \Sigma y_j^2 - n\bar{y}^2 - \Sigma x_j^2 + n\bar{x}^2.$$

Let $\hat{m}_1$ and $\hat{m}_2$ be the roots of (3). If $A \neq 0$, then

$$\hat{m}_1 = \left(B + \sqrt{B^2 + 4A^2}\right)/2A$$

and

$$\hat{m}_2 = \left(B - \sqrt{B^2 + 4A^2}\right)/2A.$$

Note that $\hat{m}_1$ has the same sign as A.

Substituting $\hat{m}_1$ and $\hat{m}_2$ into (2), one can find the corresponding values $\hat{b}_1$ and $\hat{b}_2$ of b. Thus the regression lines in (1) are determined, denote them by

(4) $$L_1: y - \hat{b}_1 - \hat{m}_1 x = 0,$$

and

(5) $$L_2: y - \hat{b}_2 - \hat{m}_2 x = 0.$$

Actually equation (3) is $[\Sigma x_j^2 - \Sigma x_j^2] \cdot m^3 + Am^2 - Bm - A = 0$.

If one views it as a cubic equation of m, then it has a third root $\hat{m}_3 = \infty$, and the corresponding regression line L_3 will be

(6) $$L_3: x - \bar{x} = 0.$$

The lines thus determined have the following properties:

Property I. All the lines L_1, L_2, and L_3 pass through the point $(\bar{x}, \bar{y})$, since $(\bar{x}, \bar{y})$ satisfies the Equations (4), (5), and (6).

Property II. L_1 and L_2 are perpendicular, since

$$\hat{m}_1 \hat{m}_2 = -A/A = -1.$$

Property III. $\Sigma d_j^2(2) \geq \Sigma d_j^2(3) \geq \Sigma d_j^2(1)$,

where $\Sigma d_j^2(k)$ is the sum of square distances from the n points $(x_1, y_1), \ldots, (x_n, y_n)$ to the line L_k $(k = 1, 2, 3)$. Equalities occur if and only if the n points coincide, or the n points are symmetric about the point $(\bar{x}, \bar{y})$ (in the latter case, n must be an even number).

Proof. First one proves that $\Sigma d_j^2(3) \geq \Sigma d_j^2(1)$. This is equivalent to showing

$$\Sigma(x_j - \bar{x})^2 - \frac{1}{1 + \hat{m}_1^2} \Sigma\left(y_j - \hat{b}_1 - \hat{m}_1 x_j\right)^2 \geq 0.$$

Now

$$\begin{aligned}
(1 + \hat{m}_1^2)\Sigma(x_j - \bar{x})^2 - \Sigma\left(y_j - \hat{b}_1 - \hat{m}_1 x_j\right)^2 & \\
&= (1 + \hat{m}_1^2)\Sigma(x_j - \bar{x})^2 - \left\{\Sigma(y_j - \hat{m}_1 x_j)^2 - n\hat{b}_1^2\right\} \\
&= -\left(\Sigma y_j^2 - n\bar{y}^2 - \Sigma x_j^2 + n\bar{x}^2\right) + 2\hat{m}_1\left(\Sigma x_j y_j - n\bar{x}\bar{y}\right) \\
&= -B + B + \sqrt{B^2 + 4A^2} \geq 0.
\end{aligned}$$

Similarly for $\Sigma d_j^2(2) \geq \Sigma d_j^2(3)$.

Property III says that the line L_1 always minimizes Σd_j^2.

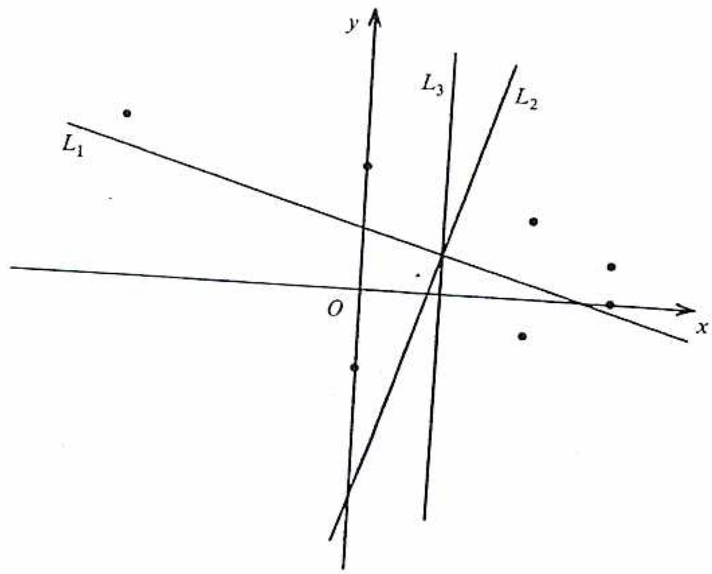
EXAMPLE. In an experiment, let seven points $(0, -2)$, $(4, -1)$, $(6, 0)$, $(6, 1)$, $(4, 2)$, $(0, 3)$ and $(-6, 4)$ be observed. Then $n = 7$, $\bar{x} = 2$, $\bar{y} = 1$, $A = -28 < 0$ and $B = -84$. From Eq. (3), one has

$$m^2 - 3m - 1 = 0,$$

$$\hat{m}_1 = (3 - \sqrt{13})/2 < 0,$$

and

$$\hat{m}_2 = (3 + \sqrt{13})/2.$$



From Eq. (2) one obtains $\hat{b}_1 = -2 + \sqrt{13}$ and $\hat{b}_2 = -2 - \sqrt{13}$,
The three lines

$$L_1: 2y - (3 - \sqrt{13})x + 2(2 - \sqrt{13}) = 0$$

$$L_2: 2y - (3 + \sqrt{13})x + 2(2 + \sqrt{13}) = 0$$

and

$$L_3: x - 2 = 0. \text{ (See figure.)}$$

The lines L_1, L_2, and L_3 give the values of Σd_j^2, 19.52, 120.48 and 112 respectively.

References

1. Wilfrid J. Dixon and Frank J. Massey, Jr., Introduction to Statistical Analysis, McGraw-Hill, New York, 1969, p. 194.

2. John E. Freund, Statistics, Prentice-Hall, New Jersey, 1970, p. 265.

THE TEACHING OF MATHEMATICS

EDITED BY MARY R. WARDROP AND ROBERT F. WARDROP

Material for this department should be sent to Professor Robert F. Wardrop, Department of Mathematics, Central Michigan University, Mount Pleasant, MI 48859.

WHY THE PRODUCT TOPOLOGY?

VLADIMIR DROBOT AND JOHN SAWKA
Department of Mathematics, University of Santa Clara, Santa Clara, CA 95053

A student is often puzzled when he or she first encounters the product topology. The usual topology on an infinite product $\Pi_\alpha X_\alpha$ of topological spaces X_α has as a basis the sets of the form $\Pi_\alpha U_\alpha$, where each U_α is open *and* $U_\alpha = X_\alpha$ for all but finitely many values of α. Why this last