PHYSICAL CHEMISTRY

Methods, Techniques, and Experiments

Rodney J. Sime

California State University,
Sacramento

SAUNDERS GOLDEN SUNBURST SERIES

SAUNDERS COLLEGE PUBLISHING

Philadelphia Ft. Worth Chicago
San Francisco Montreal Toronto
London Sydney Tokyo

Text Typeface: Times Roman
Compositor: TAPSCO, Inc.
Acquisitions Editor: John Vondeling
Managing Editor: Carol Field
Project Editor: Mary Patton
Copy Editor: Ruth Melnick
Manager of Art and Design: Carol Bleistine
Art Director: Christine Schueler
Art and Design Coordinator: Doris Bruey
Text Designer: Dorothy Chattin
Cover Designer: Lawrence R. Didona
Text Artwork: Rolin Graphics, Inc.
Director of EDP: Tim Frelick
Production Manager: Charlene Squibb

Cover Credit: Samuels Studio © 1988 Al Micciche, photographer

Printed in the United States of America

PHYSICAL CHEMISTRY: METHODS, TECHNIQUES, AND EXPERIMENTS

ISBN 0-03-009499-2

Library of Congress Catalog Card Number: 89-043114

0123 045 987654321

The Method of Least Squares

The method of least squares is a statistical method for determining the best line that represents a linear relationship between the dependent and independent variables. If there is a unique best line, then it must be characterized by a unique slope m and a unique intercept b, which are to be determined. In drawing a line by hand, we judge with our eye the relative position of the line with respect to the points, as shown for a few points in Figure 6–2. The residual r_1 for point (x_1, y_1) is the difference in the experimental value for y at x_1 and the calculated value for y at x_1, or

$$\text{residual} = y_{\text{experimental}} - y_{\text{calculated}} \tag{6–5}$$

We assume for the moment that we know the values of the slope m and the intercept b, so that we can calculate the value of y at each point and compare it with the measured value

$$r_i = y_i - (mx_i + b) \tag{6–6}$$

When we draw by hand the best line through a set of points, we use our eye to minimize the residuals. The Principle of Least Squares is quite similar. It can be shown statistically (Young, 1962) that the most probable values of x_i are those that minimize the sum of the squares of residuals. Let R be the sum of the squares of the residuals so that

$$R = \sum r_i^2 = \sum (y_i - mx_i - b)^2 \tag{6–7}$$

As usual for any function, the conditions for the minimum in the function are

$$\left(\frac{\partial R}{\partial m}\right)_b = 0 \quad \text{and} \quad \left(\frac{\partial R}{\partial b}\right)_m = 0 \tag{6–8}$$

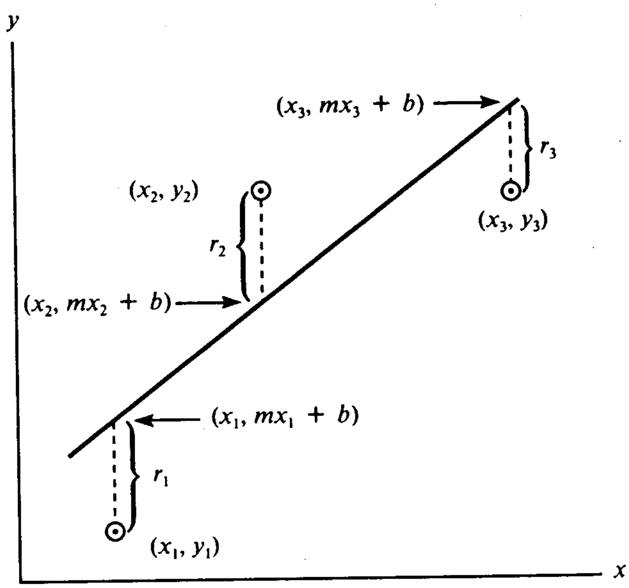

Figure 6–2 Residuals r_i.

When Equation 6–7 is expanded by squaring, the result is

$$R = \sum y_i^2 - 2m \sum x_i y_i - 2b \sum y_i + m^2 \sum x_i^2 + 2m \sum x_i b + nb^2 \quad (6\text{–}9)$$

$$\left(\frac{\partial R}{\partial m}\right)_b = 0 - 2 \sum x_i y_i - 0 + 2m \sum x_i^2 + 2b \sum x_i + 0 = 0 \quad (6\text{–}10)$$

$$\left(\frac{\partial R}{\partial b}\right)_m = 0 - 0 - 2 \sum y_i + 0 + 2m \sum x_i + 2nb = 0 \quad (6\text{–}11)$$

Equations 6–10 and 6–11 are two equations in two unknowns, m and b, which may be solved simultaneously to give the required slope and intercept of the statistically best line:

$$\text{slope} = m = \frac{n \sum y_i x_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (6\text{–}12)$$

$$\text{intercept} = b = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (6\text{–}13)$$

This is the best method of determining the equation of a line, since it is completely objective. All that is required is the calculation of the four summations that arise in Equations 6–12 and 6–13.

Now let us use the data in Table 6–1 to determine by the method of least squares the equation of the line shown in Figure 6–1 (the appropriate summations are shown in Table 6–2):

$$m = \frac{n \sum y_i x_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{6(90.125) - (12.45)(11.10)}{6(121.4125) - (12.45)^2} \quad (6\text{–}14)$$

$$= -1.184$$

TABLE 6–2

Summations for Six Data Points (n = 6) from Table 6–1

	x	y	x^2	xy
	−3.80	8.2	14.4400	−31.1600
	−1.20	6.0	1.4400	−7.2000
	0.75	3.3	0.5625	2.4750
	3.10	1.6	9.6100	4.9600
	5.60	−2.0	31.3600	−11.2000
	8.00	−6.0	64.0000	−48.0000
Σ =	12.45	11.10	121.4125	−90.1250