

Multiple Linear Least-Squares Fits with a Common Intercept: Determination of the Intrinsic Viscosity of Macromolecules in Solution

W

M. C. López Martínez, F. G. Díaz Baños, A. Ortega Retuerta, and J. García de la Torre*

Departamento de Química Física, Facultad de Química, Universidad de Murcia, 30071 Murcia, Spain; *jgt@um.es

There are a variety of situations in which sets of (x, y) data obey linear relationships with a common intercept. Such data could be analyzed in an obvious, simplistic way, carrying out an ordinary least-squares fit for each set and averaging *a posteriori* the values of the intercept. However, the values of the slopes would be influenced by the fact that the intercepts in each fit are not identical. The problem of multiple linear least-squares fits with a common intercept has a mathematically rigorous yet simple solution that has the pedagogical value of illustrating the basic principle of least-squares fitting.

Simplest Solution: Two Linear Relationships

Consider two series of data (x_i, y_i) and (x_i, z_i) , $i = 1, \dots, n$. We want to find the coefficients of the two linear equations, $y = a + b_y x$ and $z = a + b_z x$, where the slopes are b_y and b_z and the common intercept is a . The sum of square deviations, including those of the two series is,

$$\Delta = \sum_{i=1}^n \left\{ v_i [y_i - (a + b_y x_i)]^2 + w_i [z_i - (a + b_z x_i)]^2 \right\} \quad (1)$$

where v_i and w_i are the statistical weights of the (x_i, y_i) and (x_i, z_i) points, respectively. These weights can be assigned according to data statistics (as described, for instance, by Bevington, ref 1). The least-squares criterion is the minimization of Δ with respect to the coefficients,

$$\frac{\partial \Delta}{\partial a} = 0; \quad \frac{\partial \Delta}{\partial b_y} = 0; \quad \frac{\partial \Delta}{\partial b_z} = 0 \quad (2)$$

These conditions lead immediately to the following equations,

$$a \sum (v_i + w_i) + b \sum v_i x_i + c \sum w_i x_i = \sum (v_i y_i + w_i z_i) \quad (3)$$

$$a \sum v_i x_i + b \sum v_i x_i^2 = \sum v_i y_i x_i \quad (4)$$

$$a \sum w_i x_i + c \sum w_i x_i^2 = \sum w_i z_i x_i \quad (5)$$

Equations 3–5 form a system of three linear equations whose three unknowns are the coefficients a , b_y , and b_z . It is straightforward to obtain final expressions for the three coefficients by applying Cramer's rule. However, such expressions are lengthy and we have simply left the task of solving the system to our computer program.

Generalized Solution: Multiple Linear Relationships

This simple case, two sets of data points having the same abscissas, can be easily expanded to a more general situation

with s data sets, each with n_k points $(x_i^{(k)}, y_i^{(k)})$ with statistical weights $w_i^{(k)}$, where $i = 1, \dots, n_k$, $k = 1, \dots, s$. The analog of eq 1 for the sum of squares, Δ , has now s terms, each being a sum over the n_k points. The minimization condition leads to a linear system of $s + 1$ equations, the first of which is,

$$a \sum_k \sum_i w_i^{(k)} + \sum_k b^{(k)} \sum_i w_i^{(k)} x_i^{(k)} = \sum_k \sum_i w_i^{(k)} x_i^{(k)} z_i^{(k)} \quad (6)$$

and the next s equations are of the form,

$$a \sum_i w_i^{(k)} x_i^{(k)} + b^{(k)} \sum_i w_i^{(k)} [x_i^{(k)}]^2 = \sum_i w_i^{(k)} x_i^{(k)} y_i^{(k)} \quad (7)$$

which can be solved for the common intercept, a , and the slopes, $b^{(k)}$. This can be done by any standard numerical routine for solving linear equations.

Application: Charles Law

The potential applications of the multiple least-squares fit with a common intercept are numerous. An example in introductory chemistry would be the analysis of the temperature-dependent volume of a gas at constant pressure, which is governed by Charles law. A variety of laboratory setups demonstrating Charles law have been described (2, 3). Plots of temperature (Celsius) versus volume must have a fixed intercept, -273°C .

Application: Macromolecules

Molecular Weight via Light Scattering

There are other applications of this process in physical chemistry of macromolecular solutions. One example is the determination of molecular weight of the macromolecular solute using a light scattering technique (4). The intensity of scattered light depends on both solution concentration, c , and angle of observation, θ . A classical (and pedagogical) way of analyzing light scattering data is the Zimm plot (5). Teaching aids that illustrate this technique have been published (6, 7). In the Zimm plot the compound variable Kc/R , where K is an optical constant and R is the Rayleigh ratio, depends on both an angular variable q^2 [$q = (4\pi/\lambda)\sin(\theta/2)$, where λ is the wavelength of the light in solution] and concentration (5–7). Data at various c and various q are extrapolated to $c = 0$ to obtain values at zero c and variable q , and they are also extrapolated to $q = 0$ to obtain values at zero q and variable c . Standard linear least-squares fits can be used for these primary extrapolations. Thus, one obtains two sets of linear data, Kc/R versus q^2 (at zero concentration) and Kc/R versus c (at zero angle). According to theory, in the Zimm

plot these two linear graphs have a common intercept, which is $1/M$, the reciprocal of the molecular weight of the macromolecule. The determination of the linear coefficients with our method may increase the quality of the results, not only for M , but also for the slopes of both plots, which yield, respectively, the radius of gyration and the second virial coefficient.

Intrinsic Viscosity

Another example of least-squares fits with a common intercept in a classical laboratory experiment is the determination of the molecular weight of a macromolecule from the concentration dependence of solution viscosities, expressed by the intrinsic viscosity, $[\eta]$. A characteristic feature of macromolecular substances is the effect that macromolecules have on the viscosity of a solvent: when macromolecules are dissolved in a simple solvent, a minute amount of macromolecular solute produces a large increase in the viscosity of the solution with respect to that of the pure solvent. The quantity that measures this effect in dilute solution, the intrinsic viscosity $[\eta]$, is studied in undergraduate courses of physical chemistry (8), polymer science (9), molecular biophysics (10), and related fields. The intrinsic viscosity can be employed to determine the molecular mass of the macromolecule, M , from the Mark-Houwink-Sakurada equation, $[\eta] = KM^a$, whose parameters K and a are known for many polymer-solvent systems (11). This property is also informative about the details of macromolecular structure that determine the typical flexibility of most synthetic polymers (4, 12). In the case of rigid biological macromolecules the intrinsic viscosity is an indicator of their peculiar shape (13).

The determination of $[\eta]$ from measurements of solution viscosities is a simple but effective experiment that can be integrated to laboratory courses (14–16), as described by several authors in this *Journal* (17–19). In polymer science laboratories, the characterization of the polymer by its intrinsic viscosity may complement an experiment on polymer synthesis (15, 16). Accurate measurements can be achieved with low-cost equipment, provided that the procedures are carefully followed and an adequate, precise treatment of the experimental data is performed. The relative viscosity, $\eta_r = \eta/\eta_0$, is the ratio of the intrinsic viscosity of the solution, $[\eta]$, to that of the pure solvent, $[\eta_0]$. The specific viscosity is $\eta_{sp} = (\eta - \eta_0)/\eta_0 = \eta_r - 1$. The intrinsic viscosity $[\eta]$ is defined as the ratio of specific viscosity to concentration in the limit of infinite dilution:

$$[\eta] = \lim_{c \rightarrow 0} \frac{\eta_{sp}}{c} = \lim_{c \rightarrow 0} \frac{\eta - \eta_0}{\eta_0 c} = \lim_{c \rightarrow 0} \frac{\eta_r - 1}{c} \quad (8)$$

The dependence of solution viscosity on concentration, whose form is unknown in principle, can be anyway expressed as a Taylor expansion in powers of concentration, $\eta = A_0 + A_1 c + A_2 c^2 + \dots$. It is obvious that $A_0 = \eta_0$ and from eq 8, we have $A_1/\eta_0 = [\eta]$.

According to eq 8, $[\eta]$ is the intercept in a plot of η_{sp}/c versus c , and in dilute solutions this variation is linear. The equation for this linear relationship can be written as,

$$\frac{\eta_{sp}}{c} = [\eta] + k_K [\eta]^2 c \quad (9)$$

where k_H is the Huggins constant (20), defined as $k_H = A_2/(\eta_0[\eta]^2)$.

Kraemer (21) proposed to obtain the intrinsic viscosity from a different equation that is also linear in concentration and gives $[\eta]$ as the intercept,

$$\frac{\ln(\eta_r)}{c} = [\eta] - k_K [\eta]^2 c \quad (10)$$

where k_K is the Kraemer constant. This equation originates from the fact that for a very dilute solution, η_{sp} is very small, and $\ln \eta_r = \ln(1 + \eta_{sp}) \approx \eta_{sp}$. The Kraemer extrapolation has the same intercept $[\eta]$ as the Huggins equation. It can be shown (12, 13) that $k_H + k_K = 1/2$. This relationship comes from purely mathematical arguments and should hold for any polymer-solvent system.

The recommended procedure for the determination of $[\eta]$ includes the consideration of both linear relationships, in which the intercept must be the same. Therefore, our least-squares algorithm can be used to fit the two sets of data points η_{sp}/c versus c and $(\ln \eta_r)/c$ versus c with a common intercept, $[\eta]$, and slopes that yield k_H and k_K . In this case $x_i = c_i$ (the same values in the two series), $y_i = (\eta_{sp})_i/c_i$ and $z_i = (\ln \eta_r)_i/c_i$. As a computational aid for the determination of the intrinsic viscosity, we have written a computer program, VISFIT, which in addition to finding the fits with a common intercept, implements the numerical calculations of concentrations, viscosities from efflux times, et cetera. In addition to a simple mode, where viscosities are proportional to efflux times, the program works in an advanced mode that considers the calibration constants and the variation in solution density. The program accepts an input file with the primary data and produces an output file with the intermediate and final results, including $[\eta]$, k_H , and k_K . This file is ready to be imported to a graphing tool to display the results. The Huggins and Kraemer plots for a polystyrene sample in toluene at 25 °C are shown in Figure 1. The results from VISFIT are $[\eta] = 47.7 \text{ cm}^3 \text{ g}^{-1}$, $k_H = 0.347$, and $k_K = 0.151$. From the Mark-Houwink-Sakurada equation for this polymer-solvent system, $[\eta] = 0.00977 M^{0.73}$, and the molecular weight of the polystyrene sample is $M = 1.1 \times 10^5$. The quality of the results is assured by the excellent agreement of the sum of the Huggins and Kraemer constants, $k_H + k_K = 0.499$, with the exact result, $k_H + k_K = 1/2$.

In the teaching laboratory we ask our students to first make hand calculations, standard least-squares fits (with a calculator) and graphs, at least in the first simplest way based on the approximation $\eta/\eta_0 = t/t_0$ (if time allows, we even encourage the students to attempt the second procedure). Then, they repeat the data treatment using VISFIT. Thus in addition to their practical learning, the students can appreciate the utility of the program. This also gives an opportunity to teach them a less known, but yet simple application of the least-squares principle.

Supplemental Material

VISFIT for Linux and Microsoft MSDOS/Windows platform, along with sample input and output files is available. A document that describes the details of the experimental measurements and the use of VISFIT for data analysis is

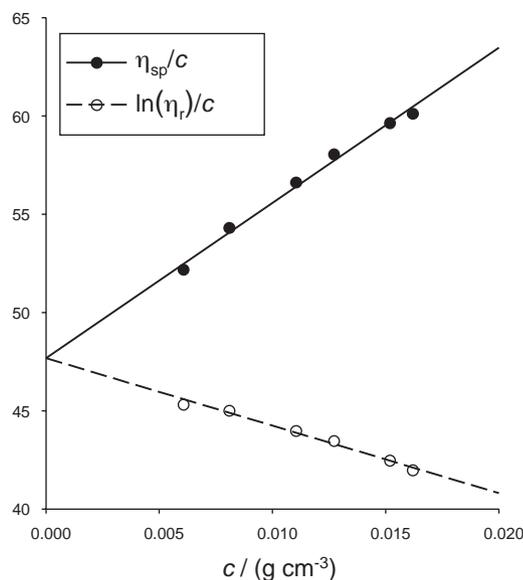


Figure 1. Plot of the double extrapolation for the intrinsic viscosity (η_{sp}/c and $\ln \eta_r/c$, in $\text{cm}^3 \text{g}^{-1}$, vs c), showing the experimental data and the extrapolating straight lines with a common intercept.

also available. This material can be found on Web site <http://leonardo.fcu.um.es/macromol> or in this issue of *JCE Online*.

Acknowledgments

This work was supported by grant PB96-1106 from the Dirección General de Investigación Científica y Técnica. Further support was provided by a grant from Fundación Séneca, ref. 2595.

Literature Cited

1. Bevington, P. R. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw Hill: New York, 1969.
2. Rose, D. *J. Chem. Educ.* **1987**, *64*, 712–713.
3. Rockley, M. G.; Rockley, N. L. *J. Chem. Educ.* **1995**, *64*, 179–181.
4. Munk, P. *Introduction to Macromolecular Science*; John Wiley & Sons: New York, 1989.
5. Zimm, B. H. *J. Chem. Phys.* **1948**, *16*, 1093–1099.
6. Poche, D. S.; Russo, P. S.; Fong, B.; Temyanko, E.; Ricks, H. *J. Chem. Educ.* **1999**, *76*, 1534–1538.
7. Matthews, G. P. *J. Chem. Educ.* **1984**, *61*, 552–554.
8. Silbey, R. J.; Alberty, R. A. *Physical Chemistry*; John Wiley & Sons: New York, 2001.
9. Billmeyer, F. W., Jr. *Textbook of Polymer Science*, 3rd ed.; Wiley Interscience: New York, 1984.
10. Van Holde, K. E.; Johnson W. C.; Ho P. S. *Principles of Physical Biochemistry*; Prentice Hall: New Jersey, 1998.
11. *Polymer Handbook*, 3rd ed.; Brandrup, J., Immergut, E. H., Eds.; John Wiley & Sons: New York, 1989.
12. Bohdanecky, M.; Kovar, J. *Viscosity of Polymer Solutions*; Elsevier: Amsterdam, 1982.
13. Harding, S. E. *Prog. Biophys. Molec. Biol.* **1997**, *68*, 207–262.
14. Daniels, F.; Alberty, R. A.; Williams, J. W.; Cornwell, C. D.; Harriman, J. E. *Experimental Physical Chemistry*, 7th ed.; McGraw-Hill: New York, 1970.
15. Sandler, S. R.; Karo, W.; Bonestel, J.; Pearce, E. *Polymer Synthesis and Characterization. A Laboratory Manual*; Academic Press: New York, 1998.
16. Collins, E. A.; Bares, J.; Billmeyer, F. W., Jr. *Experiments in Polymer Science*; Wiley Interscience: New York, 1973.
17. Mathias, L. J. *J. Chem. Educ.* **1983**, *60*, 422–424.
18. Mathias, L. J. *J. Chem. Educ.* **1983**, *60*, 990–993.
19. Richards, J. L. *J. Chem. Educ.* **1993**, *8*, 685–689.
20. Huggins, M. L. *J. Am. Chem. Soc.* **1942**, *64*, 2716–2720.
21. Kraemer, E. O. *Ind. Eng. Chem.* **1938**, *30*, 1200–1208.