NOTES

BULLETIN OF THE CHEMICAL SOCIETY OF JAPAN, VOL. 45, 2643—2644 (1972)

Application of Factor Analysis to Polarography

Darryl G. Howery

Department of Chemistry, Brooklyn College of The City University of New York, Brooklyn, N. Y., U.S.A.

(Received May 9, 1972)

Chemical problems in which the observable is measurably dependent upon several variables are seldom amenable to detailed analysis. Recent applications of the mathematical method of factor analysis to some multidimensional chemical problems have led to new insights in the fields of nuclear magnetic resonance spectroscopy[1,2] and gas-liquid chromatography.[3,4] The results of factor analyzing a data matrix of polarographic half-wave potentials for five alkaline and four alkaline-earth metals in five polar solvents are presented here.

Any function being factor analyzed must be a linear sum of factors, where in addition each factor must be in the form of a separable product term.[1] For the half-wave potential of the *i*-th ion in the α-th solvent, we assume therefore that

$$E^0_{1/2,i,\alpha} = \sum_j I_{ij} V_{j\alpha} \tag{1}$$

where I_{ij} is the *j*-th ion factor for the *i*-th ion, $V_{j\alpha}$ is the *j*-th solvent factor for the α-th solvent, and the sum is taken over the minimum number of abstract eigenvectors or controlling factors, *j*, necessary to span the data-matrix space. The reproduction portion of the factor analysis scheme allows one to find the minimum number of abstract eigenvectors, and the least-squares rotational portion of the scheme allows individual physically significant properties of both ions and solvents to be tested for possible identification with the abstract eigenvectors.

Procedure

The data matrix was taken from the compilation of Gutmann,[5] with the half-wave potentials measured relative to aqueous saturated calomel electrode being selected for factor analysis. Ions and solvents are listed in Table 1. All computations were performed on a FACOM 230-60 digital computer using the Fortran IV program developed by Malinowski and coworkers.[1]

Results and Discussion

Number of Controlling Factors. The reproduction portion of the factor analysis program enables one to reproduce the data matrix stepwise using the largest abstract eigenvector, the two largest eigenvectors, and so forth. When the agreement between the measured and the predicted half-wave potentials is overall reasonable considering the uncertainty of the measurement, the number of abstract eigenvectors used for that reproduction is taken to be equal to the number of independent controlling factors in the problem. The minimum number of factors required to adequately characterize the data space, information ordinarily not obtainable in complex problems, is thus determined. Examination of the error in the various reproductions leads one to conclude that the half-wave potential space is probably a three-factor space since with the three largest eigenvectors only seven points in the reproduced matrix are in error by more than 30 mV, the largest error being 58 mV. If the data for lithium and magnesium are ommitted from the matrix, the reproduction with three factors is not improved implying that the factors which account for irreversibility are still operative in the reduced-matrix space.

Test of Physically Significant Factors. The least-squares rotational part of the factor analysis scheme furnishes valuable insights into the factors governing the reduction process and in particular the ion-solvent interactions. A suspected chemical or physical property can be tested for identification as a fundamental factor of the space by seeing if a test vector composed of the values of that property associated with the row elements of the data matrix can be rotated mathematically onto an abstract eigenvector of the space. In a successful test the values for the property predicted by the factor analysis method must agree reasonably well with the corresponding points on the test vector. Even if some values of the property are not known or are suspect, the test can still be made since the program includes provisions for free-floating, *i.e.*, leaving blank, some points on the test vector.

A) Ion Factors: The ionic charge (see Table 1 for the results) and the reciprocal of the crystalline ionic radius both gave moderately good agreement between the test and the predicted vector. The lithium ion, the charge for which was free-floated in the ionic charge test, appears to behave more like a divalent ion. A test for the ionic potential, *i.e.*, the ionic charge divided by the crystalline ionic radius, gave the fair fit shown in Table 1. The values for Li^+ and Mg^{2+} were free-floated since the effective ionic potential quite possibly might differ from the conventional ionic potential for ions known to have high overvoltages in

---

[1] P. H. Weiner, E. R. Malinowski, and A. R. Levinstone, *J. Phys. Chem.*, **74**, 4537 (1970).

[2] P. H. Weiner and E. R. Malinowski, *ibid.*, **75**, 1207, 3160 (1971).

[3] P. H. Weiner and D. G. Howery, *Can. J. Chem.*, **50**, 448 (1972).

[4] P. H. Weiner, C. Dack, and D. G. Howery, *J. Chromatogr.*, in press.

[5] V. Gutmann, "Co-ordination Chemistry in Non-aqueous Solutions," Springer-Verlag, Vienna (1968), p. 33.

TABLE 1. TESTS FOR PHYSICALLY SIGNIFICANT PARAMETERS

Ion	Ionic charge		Ionic potential	
	Test	Predicted	Test	Predicted
Li^+	(1.00)	2.03	(1.47)	1.88
Na^+	1.00	1.16	1.03	0.92
K^+	1.00	1.25	0.75	1.02
Rb^+	1.00	0.92	0.68	0.63
Cs^+	1.00	0.87	0.60	0.57
Mg^{2+}	2.00	2.17	(3.02)	2.07
Ca^{2+}	2.00	2.06	2.02	1.94
Sr^{2+}	2.00	1.76	1.78	1.62
Ba^{2+}	2.00	1.78	1.50	1.67

Solvent	ΔH_{vap}		$1/R'$		Donor number	
	Test	Predicted	Test	Predicted	Test	Predicted
DMSO	12.6	12.3	1.47	1.45	2.98	2.72
DMA	12.7	13.5	1.45	1.45	2.78	2.82
Water	10.5	11.4	1.39	1.41	1.80	2.19
PDC	11.9	11.5	1.22	1.21	1.51	1.43
MeCN	7.8	7.8	1.22	1.21	1.41	1.30

PDC=propylene dicarbonate. Three-factor space assumed for tests. Points in parenthesis on the test vectors were free-floated. Values for the crystalline ionic radius (Å) were taken from "Handbook of Chemistry and Physics," ed. by R. C. Weast 46th Edition, Chemical Rubber Co., Cleveland (1964), p. F89; for the molar enthalpy of vaporization (kcal/mol), from reference 5 and from "Techniques of Organic Chemistry," ed. by A. Weissberger Vol. II, 3rd Edition, Wiley, New York, (1970); for the radial correction factor, from Ref. 7; and for the donor number, from Ref. 5.

water. Crystalline ionic radius, equivalent weight, solubility of the metal in mercury (g/100 g) and ionic polarizability do not appear to be factors since the fit was quite poor for each of the tests. A test for the solubility expressed as a mole fraction showed the proper trends but the fit was still rather poor.

B) Solvent Factors: Since mathematically in a three-factor space test points for all five solvents must be specified on a test vector, the lack of reported physical properties for PDC somewhat restricted the solvent properties which could be tested. Three solvent parameters tested moderately well (see Table 1 for the results): the molar enthalpy of vaporization, the donor number as defined by Gutmann,6) and the reciprocal of a radial correction factor, R', proposed by Matsuura *et al.*7) Boiling point (°K), dipole moment, molar polarization, dielectric constant, viscosity, Trouton's constant and the Hildebrand solubility parameter all gave sufficiently poor fits to be excluded as fundamental factors.

Since so little is known about the mathematical formulation of ion-solvent interactions especially when expressed in the form required by Eq. 1), speculation on the implications of the ion and solvent tests reported here seems premature. A major advantage of the method of factor analysis is that one can test an individual parameter in a many-factor space though lacking a theoretical model for the interaction.

If three independent ion factors or solvent factors are discovered, one can attempt to reproduce the data matrix using physically significant test vectors.2,4) The reproduction using the set molar enthalpy of vaporization, donor number and $1/R'$ was surprisingly good, with the average of the column average errors being 28 mV. The solvent-factor space would appear to be spanned fairly well with these three factors.

The application of factor analysis to new areas of chemistry is being severely limited by the lack of large, complete, and accurate data matrices. At present, extensive data of the type required for comprehensive factor analysis studies are available as far as we know only in the field of gas-liquid chromatography.8) The matrix employed in this study severely tested the factor analysis method since the number of solvents was so small compared to the number of factors required to span the space. Since experience has indicated that the number of rows and columns should preferably be at least twice the number of factors, the reproduction of the data matrix with three eigenvectors and the fit of the test vectors given in Table 1 are encouraging. More definitive conclusions to this complex problem should be possible if a larger data matrix were available.

Dr. K. Izutsu suggested this problem and furnished much valued advice. J. Okuda's help on programming, Professor T. Fujinaga's financing of the computer work, and the use of the Computer Center of Kyoto University are gratefully acknowledged.

---

6) Ref. 5., p. 19.

7) J. Matsuura, K. Umemoto, and Y. Kakeuchi, 16th Polarographic Symposium, Kyoto, 1970.

8) W. O. McReynolds, "Gas Chromatographic Retention Data," Preston Technical Abstracts Co., Evanston, Illinois (1966).