COMMUNICATIONS TO THE EDITOR

The Use of Dual Linear Programming in Formulating Approximating Functions by Using the Chebyshev Criterion

P. K. LEUNG
Imperial Oil Limited, Sarnia, Ontario, Canada
and DONALD QUON
University of Alberta, Edmonton, Alberta, Canada

Suppose that we are trying to establish a functional relationship between a dependent variable t and the independent variables x, y, and z, given a set of m values of t_i for $x = x_i$, $y = y_i$, $z = z_i$ ($i = 1, 2, \ldots, m$). There are many reasons why it is convenient to have the approximate function $f(x, y, z)$ in the following form:

$$f(x, y, z) = \sum_{j=0}^{n} \alpha_j \, g_j(x, y, z) \qquad (1)$$

where $g_j(x, y, z)$ are specified functions and α_j are the expansion coefficients to be determined. It is assumed that physical insight, intuition, and experience, etc., would be brought to bear in the selection of the specified functions.

The differences between the values predicted by Equation (1) and the observed or known values of t_i

$$\epsilon_j = t_i - f(x_i y_i z_i) \qquad i = 1, 2, \ldots, m$$

may be used to establish some criterion of best fit.

The least squares criterion, which minimizes $\sum_{i=1}^{m} \epsilon_i^2$, is generally conceded to be the proper one to use for correlating data whose errors follow a Gaussian distribution. However, there are many cases where the Chebyshev criterion, which minimizes the maximum error (min.max $|\epsilon_i|$ for $i = 1, 2, \ldots, m$) may be more appropriate; for example, if an approximating function is used to represent known or smoothed data, then an upper bound on the error is often desirable. Furthermore, there are other cases where computational difficulties associated with the least squares approach may make it unacceptable.

Algorithms for determining the expansion coefficients with the use of the Chebyshev criterion are available (*1*). These employ a heuristic approach and are iterative in nature. However, we can also pose the problem as a linear programming problem, and this approach has many attractive features. The method has built-in tests for existence and uniqueness of solutions, and it always converges.

MATHEMATICAL DEVELOPMENT

The problem may be stated as follows:

Given a set of m real functional values t_i at $x = x_i$, $y = y_i$, $z = z_i$ ($i = 1, 2, \ldots, m$) and n specified functions $g_j(x, y, z)$, ($j = 1, 2, \ldots, n$) continuous over a closed region R, find a n dimensional real vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$ such that λ, defined below, is a minimum.

$$\left| t_i - \sum_{j=1}^{n} \alpha_j \, g_j(x_i y_i z_i) \right| \leq \lambda \qquad i = 1, 2, \ldots, m \qquad (2)$$

The problem may be restated as the following linear programming problem:

Minimize

$$Z = \lambda \qquad (3)$$

Subject to

$$\lambda + \sum_{j=1}^{n} \alpha_j g_j(x_i y_i z_i) \geq t_i \qquad (4a)$$

$$\lambda - \sum_{j=1}^{n} \alpha_j g_j(x_i y_i z_i) \geq -t_i \qquad i = 1, 2, \ldots, m \qquad (4b)$$

It is readily seen that (4*a*) and (4*b*) are equivalent to (2).

The above problem can now be solved by the standard simplex method, developed by Dantzig (2). However, the number of constraint equations is $2m$ and if m, the number of data points, is large, then the system may be too large for practical computation. The difficulty may be overcome by solving the dual problem, as described by Wagner (3).

By definition, if two linear programming problems are said to be dual, they are related to one another as follows:

When the minimizing problem is considered as the primal, that is

Minimize

$$Z = c_1\alpha_1 + c_2\alpha_2 + \dots c_n\alpha_n \tag{5}$$

Subject to

$$\sum_{j=1}^{n} a_{ij}\alpha_j \geqq b_i \quad i = 1, 2, \dots m \qquad \text{all } \alpha_j \geqq 0 \tag{6}$$

then the following maximizing problem is called the dual.

Maximize

$$W = b_1\beta_1 + b_2\beta_2 + \dots b_m\beta_m \tag{7}$$

Subject to

$$\sum_{i=1}^{m} a_{ji}\,\beta_i \leqq c_j \quad j = 1, 2, \dots n \qquad \text{all } \beta_i \geqq 0 \tag{8}$$

It has been shown (4) that:

1. In the final solution, when the objective function of the primal is minimized, that of the dual is maximized at the same point, that is

$$Z_{\min} = W_{\max} \tag{9}$$

2. The coefficient matrix of one system is the transpose of the other.

3. The constants on the right-hand side of one system are the coefficients of the objective function of the other.

As a general case, consider the primal to be a mixed system consisting of the following constraints:

k constraints of type

$$\Sigma\ a_{ij}x_i = b_j \quad j = 1, \dots, k$$

r constraints of type

$$\Sigma\ a_{ij}x_i \leqq b_j \quad j = k + 1, \dots, k + r$$

s constraints of type

$$\Sigma\ a_{ij}x_i \geqq b_j \quad j = k + r + 1, \dots, k + r + s$$

Some of the variables are non-negative, while the rest are unrestricted in sign. This primal system is the most general system possible and is represented diagrammatically in Table 1 with A, B, C, D, E, and F denoting the subarrays of the coefficient matrix.

Before dualizing the primal, it is necessary to multiply the constraints $k + 1$, $k + 2$, . . ., $k + r$ by -1 so as to reverse the inequalities. The dual is shown schematically in the lower diagram of Table 1, with the superscript T denoting the transpose of the original subarray of the coefficient matrix.

Using Table 1 as a guide, we see that the primal problem posed by Equations (3) and (4) gives rise to the following dual problem:

Maximize

$$W = \sum_{i=1}^{m} t_i\,(\psi_i - \phi_i) \tag{10}$$

Subject to

$$\sum_{i=1}^{m} (\psi_i + \phi_i) \leqq 1 \tag{11}$$

and

$$\sum_{i=1}^{m} (\psi_i - \phi_i) \sum_{j=1}^{n} g_j\,(x_i y_i z_i) = 0 \tag{12}$$

To convert constraint (11) into an equation, a slack variable has to be introduced. However, if we assume that we are dealing with the nontrivial case in which $\lambda > 0$, then this slack variable cannot be in the final basis, and must therefore be zero. Thus constraint (11) becomes

TABLE 1. DUALITY OF LINEAR PROGRAMMING PROBLEMS

Primal

	x_1 Non-negative x_p	x_{p+1} Unrestricted x_n		
1 … k	A	B	$=$	b_1 … b_k
$k+1$ … $k+r$	C	D	$\leqq$	b_{k+1} … b_{k+r}
$k+r+1$ … $k+r+s$	E	F	$\geqq$	b_{k+r+1} … b_{k+r+s}
	c_1 … c_p	c_{p+1} … c_n		min

Dual

	Unrestricted y_1 … y_k	y_{k+1} … y_{k+r}	Non-negative y_{k+r+1} … y_{k+r+s}		
1 … p	A^T	$-C^T$	E^T	$\leqq$	c_1 … c_p
$p+1$ … p_n	B^T	$-D^T$	F^T	$=$	c_{p+1} … c_n
	b_1 … b_k	$-b_{k+1}$ … $-b_{k+r}$	b_{k+r+1} … b_{k+r+s}		max

$$\sum_{i=1}^{m} (\psi_i + \phi_i) = 1 \quad (13)$$

The optimum basic vector of the dual system will correspond to those constraints in the primal which have a zero value for the slack variable, that is, those for which the error ϵ_i is in fact equal to λ. This is in accordance with the theorem of Poussin (5), which states that the solution vector α which minimizes λ gives a subset of n points (in the original set of m points) which will have an error equal to λ. Since these particular constraints become equalities, the coefficients α_j and λ can be easily calculated.

This optimum basic vector also gives the distribution of the points in the region giving the worst fit. If these points are congregated about a point, this suggests a large random error for that point; thus the point should be eliminated from the system and the analysis repeated.

The advantage of this approach is that the number of constraints in the dual is only $(n + 1)$—corresponding to the n coefficients plus one for λ—compared with $2m$ constraints for the primal. Since the number of iterations required to reach the optimum solution depends primarily on the number of constraint equations, the saving in computational effort is obvious.

NUMERICAL EXAMPLE

The coefficients of the Benedict-Webb-Rubin equation of state were determined numerically by using various criteria of best fit. Explicit expressions for calculating z, the compressibility factor, and for $(H - H^o)_T$, the enthalpy departure (6), were compared to the experimental values of these quantities for propane. The data included two hundred and thirty-seven smoothed points from the volumetric data compiled from Deschner and Brown (7) and eleven points on enthalpy departure by Yarborough and Edmister (8).

The calculations were straightforward, except that γ, which occurs in the exponential term, had to be determined by trial and error. A search technique was used to find the optimal γ, that is, that value which gave the best fit for the whole set of coefficients.

Four cases were investigated and the results given in Table 2:

Case I. Least squares criterion with standard techniques (9) and the volumetric data only.

Case II. Chebyshev criterion, minimizing the maximum percent deviation in z and using the volumetric data only.

Case III. Chebyshev criterion, minimizing the maximum absolute deviation in z and using the volumetric data only.

In Cases II and III, as compared to Case I, there was a significant decrease in the maximum errors but there was some increase in the average error. Which is more desirable depends obviously on the problem.

Case IV. Chebyshev criterion with constraints. Here an attempt was made to fit both the volumetric data and the enthalpy departure data simultaneously. The Chebyshev criterion, minimizing the maximum error in z, was used together with the constraint that the error in predicting enthalpy departure must not exceed 2.0 B.t.u./lb. It should be noted upon examining case I in Table 2 that the least squares fit is excellent for the volumetric data but not at all good for the enthalpy departure data. This is not surprising, since calculation of the latter involves derivatives, and it is well known that the prediction of derivative values by numerical methods is hazardous. However, incorporation of enthalpy departure data into the fit by means of the constraint equations produces a set of coefficients (case IV in Table 2) which will satisfactorily represent all the data. Of the four cases studied, this is probably the most useful for the design engineer.

TABLE 2. CLOSENESS OF FIT USING VARIOUS COEFFICIENTS FOR B-W-R EQUATION* FOR PROPANE

Case	I	II	III	IV
$A_o \times 10^{-4}$	2.5981	0.8151	0.8847	0.7747
B_o	1.4833	0.1710	0.1213	0.1842
$C_o \times 10^{-9}$	5.6939	9.5383	9.1613	9.6362
$a \times 10^{-4}$	4.0378	6.8293	11.381	7.8955
b	4.7480	6.6017	10.427	7.3152
$a\alpha \times 10^{-5}$	1.1230	1.3186	1.6274	1.3972
$c \times 10^{-10}$	2.0117	2.6289	3.4856	2.8185
γ	5.6536	4.80	4.51	4.60
Ave. dev. z	0.0035	0.0053	0.0058	0.0051
Max. dev. z	0.0188	0.0200	0.0138	0.0141
Max. % dev. z	4.441	3.414	3.893	4.144
Max. dev. $(H - H^\circ)_T$	7.08	3.06	3.89	2.0

Case I: Least squares criterion

Case II: Chebyshev criterion, minimizing the maximum % deviation in z

Case III: Chebyshev criterion, minimizing the maximum absolute deviation in z

Case IV: Chebyshev criterion with constraints, minimizing the maximum absolute deviation in z with the constraint that the maximum permissible deviation in the enthalpy departure predication is 2.0 B.t.u./lb.

* English units are used.

It should be further noted that the four sets of coefficients are markedly different from each other; in any empirical fit it would be difficult to assign a physical significance to the value of the coefficients.

CONCLUSION

There are two areas in formulating approximating functions where linear programming techniques should have wide applicability: (1) where the nature of the problem indicates that the Chebyshev criterion of best fit is the proper one; and (2) where the conventional least squares approach leads to computational difficulties due to (a) degeneracy in the normal set of equations, or (b) the necessity for incorporating inequality constraints on the system of equations, which could arise from information on derivatives, integral equations, etc. For either (a) or (b) dual linear programming technique provides a feasible, alternative curve-fitting procedure.

ACKNOWLEDGMENT

Dr. P. K. Leung, while working on this project, was supported by a National Research Council Studentship. R. Clare, a graduate student, obtained the numerical results presented.

LITERATURE CITED

1. Ralston, A., "A First Course in Numerical Analysis," p. 301, McGraw-Hill, New York (1965).
2. Dantzig, G., "Linear Programming and Extensions," Chap. 5, Princeton Univ. Press, N. J. (1963).
3. Wagner, H. W., *J. Am. Stat. Assn.*, 309 (March, 1963).
4. Dantzig, G., *ibid.*, Chap. 6.
5. Poussin, C., J. de la Vallie, Pt. 2, *Ann. Soc. Sci. Bruxelles*, Memoires, **35**, 1 (1911); translated with a commentary by H. E. Salzer.
6. Benedict, M., G. B. Webb, and L. C. Rubin, *Chem. Eng. Progr.*, **47**, No. 8, 419 (1951).
7. Deschner, W. W., and G. G. Brown, *Ind. Eng. Chem.*, **32**, 836 (1940).
8. Yarborough, L., and W. C. Edmister, *A.I.Ch.E. J.*, **11**, No. 3, 492 (1965).
9. Lapidus, Leon, "Digital Computation for Chemical Engineers," p. 323, McGraw-Hill, New York (1962).