CONCLUSIONS

Profile position control appears to offer a simple, practical solution to the problem of controlling distillation columns with sharp temperature profiles.

ACKNOWLEDGMENTS

This work was done while the author was employed at the Engineering Department of E. I. du Pont de Nemours & Company during the summer of 1970. The comments and suggestions of P. S. Buckley and C. A. Loflin are gratefully acknowledged.

NOTATION

B = bottoms product flow, mole/min.
D = overhead product flow, mole/min.
F = feed rate, mole/min.
FC = flow control loop
LC = level control loop
$T5$ = temperature on tray 5, °F.
TT = temperature transmitter
TC = temperature controller
X_B = composition bottoms product, mole fraction more volatile component
X_D = composition overhead product, mole fraction more volatile component
V = vapor boilup, mole/min.
Z = profile position, number of plates from base

LITERATURE CITED

1. Luyben, W. L., *Ind. Eng. Chem. Fundamentals,* **8,** 739 (1969).
2. Luyben, W. L., *AIChE J.*, 17, 713 (1971).
3. Wilson, J. H., *Instrum. Technol.*, **18**, 37 (1971).

On Optimality Criteria for Constrained Optima

G. V. REKLAITIS

School of Chemical Engineering
Purdue University, Lafayette, Indiana 47907

In a recent paper, Law and Fariss (*9*) introduced the generalized matrix inverse to aid in the verification of the well-known Lagrangian first-order necessary and, especially, the second-order sufficient conditions for equality constrained optima. Relying on the inverse construction, they sought, furthermore, to extend both classical results to accommodate redundant constraints and the sufficient conditions to encompass inequality constraints. The purpose of this note is threefold: 1. to indicate that the proposed extension of the necessary conditions is not universally valid; 2. to demonstrate that the claimed extension to inequality constraints cannot be made and to provide a correct statement of that result; 3. to correct some inaccuracies in the use of the generalized inverse constructions.

In the notation of (*9*) which will be used throughout the subsequent discussion the problem under consideration is to identify the optimum values of a function $q(x)$ of the n-vector argument x subject to constraints $f_j(x) = 0$, $j = 1, \ldots, m$.

Necessary Conditions. The necessary conditions suggested by Law and Fariss are that there exist multipliers λ, one for each constraint such that

$$g + J^T\lambda = 0$$

where g is the gradient of q and J is the Jacobian of f, with *no restrictions* on the rank of J. The classical result requires J to have *full row rank* (*3*). As illustrated by the following well-known example (*8*) unqualified relaxation of the rank restriction leads to difficulties.

Example 1:

maximize x_1

subject to $x_2 + (x_1 - 1)^3 \leq 0$ and $x_2 \geq 0$

A graph of the feasible region readily indicates the solution (1, 0). Rewritten in equality form by the introduction of slack variables x_3 and x_4, the problem becomes

maximize x_1

subject to $x_2 + (x_1 - 1)^3 + x_3^2 = 0$

$x_2 - x_4^2 = 0$

Following the authors' relaxed form of the necessary conditions, there must exist at the point $x^0 = (1, 0, 0, 0)$ multipliers λ_1 and λ_2 such that

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = 0$$

Obviously, no such solution exists. Hence, the contradiction is forced that x^0 is not a local maximum. The example, thus, invalidates the proposed relaxed Lagrangian conditions. However, other examples presented in (*9*) indicate that there do exist situations in which the extension is valid. The problem, then, which the authors have disregarded, is to provide a general criterion for establishing whether a test point at which the constraint gradients are linearly dependent is of a benign type (such as occurs when the constraints have been unwittingly duplicated in problem formulation) in which case the relaxed condition is valid, or if it is of a pathological type (such as example 1) in which case the relaxed condition is invalid. This problem has been actively discussed in the optimization literature (*1* to *8*, *10*, *11*). Numerous abstract regularity conditions—constraint qualifications have been formulated but, especially in the equality constrained case, no computationally satisfactory approach has yet appeared (*11*). An alternative approach which obviates the need for regularity conditions is available; however, in the equality constrained case, second-order constraint information is required (*14*).

Sufficient Conditions. In essence if not in precise form Law and Fariss state the following sufficient condition for equality constrained extrema:

If the functions q and f_j are twice continuously differentiable, if there exists λ^* such that x^* and λ^* satisfy

$$g + J^T\lambda^* = 0 \qquad \text{(i)}$$

and if for every nonzero vector z such that $Jz = 0$ it follows that

$$z^T H_L z > 0 \quad (< 0) \qquad \text{(ii)}$$

then x^* is a local minimum (maximum) of the equality constrained problem.

It is further claimed that the above result is valid for inequality constrained problems "in that active inequalities can be treated simply as equalities for purposes of the sufficiency test."

Even allowing for a correct statement of the linear conditions applicable in the inequality case, this assertion is inaccurate.

Example 2:

$$\text{Extremize } \exp(-x_2) + x_1^3$$

$$\text{subject to } x_1^2 + x_2 \geqq 0$$

$$-x_2 - x_3^2 \geqq 0 \text{ and } x_3 \leqq 0$$

At $x^0 = (0, 0, 0)$, the constraints are all active. The linear conditions (i) are satisfied for any choice $\lambda_3 = 0$ and $\lambda_1 - 1 = \lambda_2$. Condition (ii) requires that

$$z^T \left\{ \begin{bmatrix} 0 & & \\ & 1 & \\ & & 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 & & \\ & 0 & \\ & & 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 & & \\ & 0 & \\ & & -2 \end{bmatrix} + \lambda_3(0) \right\} z > 0 \quad (< 0)$$

for all z satisfying

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} z = 0$$

This simplifies to the requirement that

$$z_1^T (2\lambda_1) z_1 = 2\lambda_1 z_1^2 > 0 \quad \text{for all } z_1.$$

Hence, for any choice $\lambda_1 > 0$ or $\lambda_1 < 0$ (the correct first-order necessary conditions applicable in the inequality constrained case actually require $\lambda_1 \leqq 0$) the sufficient conditions are satisfied. In the former case, the conclusion, x^0 is a local minimum, is forced; in the latter, x^0 is a local maximum. Neither case is true: for small $\epsilon > 0$, the choice $x_2 = -\epsilon^2$, $x_1 = -\epsilon^{1/3}$ yields that $q(\epsilon) \simeq 1 - \epsilon < 1 = q(0)$; while the feasible choice $x_1 = \epsilon > 0$, $x_2 = x_3 = 0$, yields that $q(\epsilon) = 1 + \epsilon^3 > 1 = q(0)$. Evidently the sufficiency test given in (9) can not be extended as claimed. Instead, the sufficient conditions for the inequality constrained problem should be stated as follows (*7*):

If the functions q and f_j are twice continuously differentiable and if there exist multipliers λ^* such that x^*, λ^* satisfy 1. $f_j(x^*) \geqq 0$; 2. $\lambda^*_j f_j(x^*) = 0$; 3. $\lambda^*_j \leqq 0$; and 4. $g + J^T\lambda^* = 0$; and if for every nonzero z such that $\nabla f_j(x^*)z = 0$, for all j in $\{j:\lambda_j < 0\}$, it follows that 5. $z^T H_L z > 0$ (< 0); then x^* is a local minimum (maximum) of the inequality constrained optimization problem.

Applied to the above example, conditions 1 through 4 are satisfied for $\lambda_3 = 0$, $\lambda_1 \leqq 0$, and $\lambda_2 = \lambda_1 - 1$. Condition 5 is not satisfied since for all z such that $\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} z = 0$, that is, all z with $z_2 = 0$, the quadratic form

$$(z_1, z_3) \begin{bmatrix} 2\lambda_1 & \\ & 2(1-\lambda_1) \end{bmatrix} \begin{bmatrix} z_1 \\ z_3 \end{bmatrix}$$

is indefinite for choice of $\lambda_1 < 0$ and semidefinite for $\lambda_1 = 0$.

Note that the sufficient conditions in either the equality or inequality constrained cases do not require a constraint qualification.

GENERALIZED INVERSE CONSTRUCTIONS

Law and Fariss state the following computational form of the sufficient conditions formulated in terms of the normalized generalized inverse.

At a given point x^*, let

$$\lambda^* = -(J^n)^T g \qquad (1)$$

and let H_L be the Hessian of the Lagrangian function evaluated at x^* with λ^* as given above. If $T^T{}_{n\times p}\, H_L\, T_{n\times p}$ is positive definite (negative definite) then x^* is a local minimum (maximum).

Here J^n denotes the generalized inverse of J and T is a transformation matrix described in (9).

Example 3:

$$\text{minimize} \quad x_1^2 - x_2x_3 + 2x_3^2$$

$$\text{subject to} \quad x_1^2 + x_2^2 = 1$$

$$x_1 - x_3^3 = 1$$

at the point $x^0 = (1, 0, 0)$, $g = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$ and $J = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

Following the computation outlined in (9)

$$T = \left[\begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] = (T_{n\times r} \mid T_{n\times p})$$

$$J^n = \begin{bmatrix} 2/5 & 1/5 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\lambda = -(J^n)^T g = \begin{bmatrix} -4/5 \\ -2/5 \end{bmatrix}$$

Evaluated with λ

$$H_L = \begin{bmatrix} 2/5 & 0 & 0 \\ 0 & -8/5 & -1 \\ 0 & -1 & 4 \end{bmatrix}$$

Finally

$$T^T{}_{n\times p}\, H_L\, T_{n\times p} = \begin{bmatrix} -8/5 & -1 \\ -1 & 4 \end{bmatrix}$$

The matrix is clearly *negative* definite—indicating a *maximum* at x^0. Yet, for any choice $\epsilon > 0$ sufficiently small, with $x_1 = 1 - \epsilon$, $x_2 = +(2\epsilon - \epsilon^2)^{1/2}$ and $x_3 = -\epsilon^{1/3}$

$$q(\epsilon) \simeq 1 + 2\epsilon^{2/3} > q(x^0) = 1.$$

A correct statement of the sufficient conditions formulated in terms of the generalized inverse should be as follows:

At a given point x^*, if $(J^nJ)g = g$ $\qquad$ (2)

let

$$\lambda^*(y) = -(J^n)^T g + (I - JJ^n)y \qquad (3)$$

Then, if there exists y^* such that $T^T{}_{n\times p}\, H_L\, T_{n\times p}$ is positive definite (negative definite) then x^* is a local minimum (maximum).

Note, H_L is evaluated at x^* and with $\lambda^* = \lambda^*(y^*)$.

Testing these conditions with Example 3, we obtain

$$J^n J\, g = \begin{bmatrix} 2/5 & 1/5 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = g$$

Then

$$\lambda = \begin{bmatrix} -4/5 \\ -2/5 \end{bmatrix} + \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

or since y is arbitrary and columns of the 2×2 matrix dependent

$$\lambda = \begin{bmatrix} -4/5 \\ -2/5 \end{bmatrix} + \begin{bmatrix} 1 \\ -2 \end{bmatrix} y_1$$

For the choice of $y_1 = 1$, $\lambda = \begin{bmatrix} 1/5 \\ -12/5 \end{bmatrix}$ and

$$T^T_{n\times p} H_L T_{n\times p} = \begin{bmatrix} 2/5 & -1 \\ -1 & 4 \end{bmatrix}$$

The matrix is clearly *positive* definite, indicating a *minimum* at x^0, as required.

The corrected sufficient conditions are based upon the following well-known result stated in terms of the normalized generalized inverse (*12*).

Theorem: The linear system $Ax = b$ has a solution if and only if

$$AA^n b = b \qquad (4)$$

and, if a solution exists, all are given by

$$x = A^n b + (I - A^n A) y \qquad (5)$$

where y is arbitrary.

The Lagrangian Necessary Conditions which from the first part of the sufficient conditions require the existence of a solution λ to the system

$$g + J^T \lambda = 0$$

Using the preceding result, the Necessary Condition can be reformulated to the following

If x^* is an extremum then

$$(J^T)(J^n)^T g = g \qquad (6)$$

and if this condition is satisfied then a set of multipliers is given by

$$\lambda = -(J^n)^T g + (I - (J^n)^T (J)^T) y \qquad (7)$$

with y arbitrary.

Note, Equation (6) is the actual necessary condition which must be verified at a given point. [See (*15, 16*) for applications of this condition to algorithms.] If Condition (6) is satisfied, then the calculation of λ via Equation (7) is meaningful. By itself Equation (7) is useless, since it will yield a value of λ for any arbitrary J and g.

Conditions (6) and (7) are equivalent to Equations (2) and (3) above if appropriate symmetry properties of the generalized inverse are taken into account. The remainder of the sufficient condition follows through as described in (9) and will not be repeated here.

Law and Fariss in stating their conditions inexplicably dropped Condition (2) and the second term of Condition (3). By dropping this second term they force the necessary conditions to yield a unique value of λ. This need not be the case when the constraint gradients are linearly dependent (see Example 2, for instance) and accounts for the failure of their condition when tested against Example 3.

The corrected version of the transformed sufficient conditions does appear to have some advantage over the conventional formulation for cases in which constraint gradient dependencies occur This advantage stems in part from the fact that only $m - r$ parameters y, rather than all m parameters λ, need to be assigned values in checking the conditions. In the full row rank case it is not clear whether the generalized inverse version represents a computational savings over existing criteria, particularly (*13*) which in general does not require the complete Hessian.

In summary, it is found that the necessary conditions advanced by Law and Fariss are valid only under certain regularity conditions. In the absence of any weaker condition offered by the authors, the only conveniently verifiable one remains the requirement of full row rank of the Jacobian. The transformed sufficient conditions are valid only for equality constrained problems and then only in the corrected form presented in this note.

NOTATION

I = the identity matrix of appropriate dimension
J = $m \times n$ = Jacobian matrix, $J = (\partial f_j/\partial x_i)$
g = $n \times 1$ = gradient vector of q
H_L = the $n \times n$ Hessian matrix of second partial derivatives of the Lagrangian function
m = number of constraints
n = number of variables
p = $n - r$
$q(x)$ = the objective function
r = the rank of J
T = $n \times n$ transformation matrix
$T_{n\times r}$, $T_{n\times p}$ = the first r, p columns of T respectively
x = $n \times 1$ vector of variables
λ = $m \times 1$ vector of Lagrange multipliers

LITERATURE CITED

1. Abadie, J., in J. Abadie (ed.), "Nonlinear Programming," pp. 21-36, North Holland, Amsterdam, Neth. (1967).
2. Arrow, K. J., L. Hurwicz, and H. Uzawa, *Nav. Res. Logistics Q.*, **8**, 175 (1961).
3. Beightler, C., and D. J. Wilde, "Foundations of Optimization Theory," Prentice-Hall, Englewood Cliffs, N. J. (1967).
4. Canon, M., C. Cullum, and E. Polak, *Soc. Ind. Appl. Math. on Control*, **4**, 528 (1966).
5. Cottle, R. W., Research Mem. RM-3858-RR, Rand Corporation, Santa Monica, Calif. (1963).
6. John, F., in "Studies and Essays: Courant Anniversary Volume," pp. 187-204, Interscience, New York (1948).
7. King, R. P., *Ind. Eng. Chem. Fundamentals*, **3**, 484 (1966).
8. Kuhn, H. W., and A. W. Tucker, *Proc. Second Berkeley Symposium Math. Stat. Prob.*, 481 (1951).
9. Law, Victor J., and Robert H. Fariss, *AIChE J.*, **17**, 425 (1971).
10. Mangasarian, O. L., and S. Fromovitz, *J. Math. Anal. Appl.*, **17**, 37 (1967).
11. Mangasarian, O. L., "Nonlinear Programming," McGraw-Hill, New York (1969).
12. Penrose, R., *Proc. Cambridge Phil. Soc.*, **51**, 406 (1955).
13. Reklaitis, G. V., and D. J. Wilde, *Operations Research*, **17**, 425 (1969).
14. Reklaitis, G. V., and D. J. Wilde, in "Optimizing Methods in Statistics," Academic Press, in press.
15. Rosen, J. R, *J. Soc. Ind. Appl. Math.*, **8**, 181 (1960).
16. Fletcher, R., in J. Abadie (ed.), "Integer and Nonlinear Programming," Ch. 6, North Holland, Amsterdam, Neth. (1970).