† *This paper was supported by the Office of Naval Research, Contract No. N00014-67-A-0321-0003 (NR047-095).*

[1] *Graduate School of Business Administration and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

[2] *Department of Statistics and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.*

# ON USING EQUALITY-CONSTRAINT ALGORITHMS FOR INEQUALITY CONSTRAINED PROBLEMS†

J. P. Evans[1] *and* F. J. Gould[2]

*Department of Statistics*
*University of North Carolina at Chapel Hill*

**Institute of Statistics Mimeo Series No. 776**

October, 1971

ON USING EQUALITY-CONSTRAINT ALGORITHMS

FOR INEQUALITY CONSTRAINED PROBLEMS[†]

by

J. P. Evans[1] and F. J. Gould[2]

Consider the inequality constrained problem

$P_0$: $$\max\ f(x), \text{ subject to}$$
$$g_i(x) \le 0,\ i = 1,\dots,m$$

where all functions are real valued and continuous on $R^n$. Suppose $x^*$ is an optimal solution to $P_0$ and define $I = \{i : g_i(x^*) = 0\}$. Recently a number of references have appeared (see, for example, [2,4]) which discuss the possibility that certain techniques for solving equality constrained problems can be effectively extended to problems like $P_0$ by first identifying I and then solving the problem

$\hat{P}$: $$\max\ f(x), \text{ subject to}$$
$$g_i(x) = 0,\ i \in I.$$

There has also been some evidence of computational success in implementing such techniques and solving an assortment of "standard" test problems [3]. The purpose of this note is to establish conditions on problem $P_0$ which make its replacement by $\hat{P}$ valid, but also to present theoretic difficulties which must inevitably be reckoned with by those researchers who would wish to elaborate the development of these techniques. The difficulties

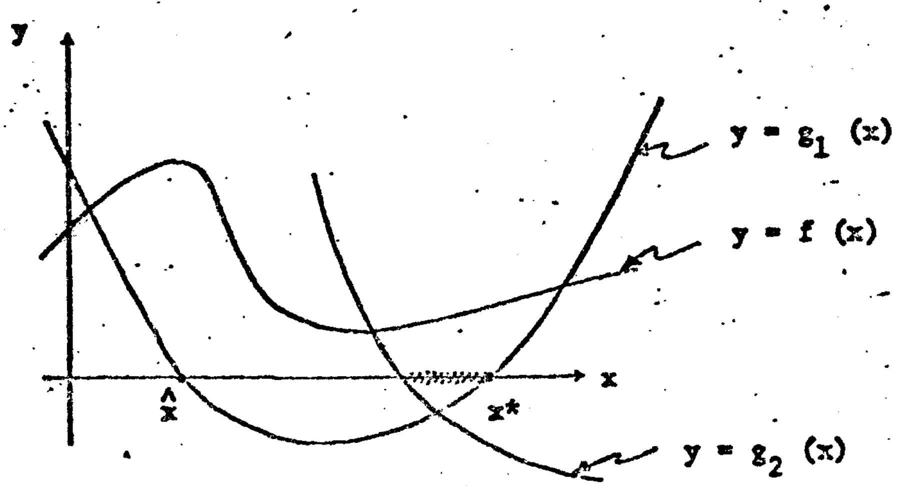
---

[†]This work was supported by the Office of Naval Research, Contract No. N00014-67-A-0321-0003 (NR047-095).

[1]Graduate School of Business Administration and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.

[2]Department of Statistics and Curriculum in Operations Research and Systems Analysis, University of North Carolina at Chapel Hill.

we refer to are due mainly to the facts that (i) a concave problem $P_0$ in general does not produce a convex constraint set in $\hat{P}$, and (ii) known penalty functions for $\hat{P}$, in the case when f is concave and the constraint functions convex, are not themselves concave functions. This point will be further discussed below. Prior to that, we first wish to note that there appears to be a not uncommon impression that $P_0$, at least in theory, can always be replaced with $\hat{P}$. The following example shows clearly that this is generally not true.




Here the constraint set $S_0 = \{x : g_i(x) \leq 0,\ i = 1, 2\}$ is shaded, $x^*$ solves $P_0$, $I = \{1\}$, and $\hat{x}$ solves $\hat{P}$. The following lemma gives a set of conditions under which, in theory, the replacement of $P_0$ with P is valid.

Lemma: Suppose $x^*$ is a global solution to $P_0$, f is strictly quasi-concave[1]

---

[1] f is strictly quasi-concave on $R^n$ if for each pair of points $x^1, x^2 \in R^n$, $f(x^1) > f(x^2)$ implies $f(tx^1 + (1 - t)\,x^2) > f(x^2)$, $0 < t < 1$.

on $R^n$, and the set $S_i = \{x: g_i(x) \leq 0\}$ is convex for each $i \in I$.[1] Then $x^*$ is a global solution to $\hat{P}$.

Proof: Define $S_0^I = \{x: g_i(x) \leq 0, i \in I\}$; suppose $\hat{x} \in S_0^I$ but $f(\hat{x}) > f(x^*)$. Define $x_t = t\hat{x} + (1-t)x^*$, $0 \leq t \leq 1$. Suppose $t > 0$; by strict quasiconcavity of $f$, $f(x_t) > f(x^*)$. Thus since $x^*$ solves $P_0$, it follows that $x_t \notin S_0$, $0 < t \leq 1$.

Since $S_i$ is convex for each $i \in I$, and $\hat{x}$, $x^* \in S_0^I$, the line segment connecting $x^*$ and $\hat{x}$ is contained in $S_0^I$. But we will show that there is a $t \in (0,1]$ such that $x_t \notin S_0^I$, yielding a contradiction. For $i \notin I$, $g_i(x^*) < 0$. By continuity of $g_i$'s, there is a $t_1 \in (0,1]$ such that $0 < t \leq t_1$ implies $g_i(x_t) < 0$, $i \notin I$. Now since $t \in (0,1]$ implies $x_t \notin S_0$, it must be that for some fixed $t \in (0, t_1]$, there is an index $i_0 \in I$ such that $g_{i_0}(x_t) > 0$. Thus, $x_t \notin S_0^I$. This contradiction shows that $\hat{x} \in S_0^I$ implies $f(\hat{x}) \leq f(x^*)$. Thus, $x^*$ is a global solution to $\max f(x)$, subject to $x \in S_0^I$, and hence a global solution to $\hat{P}$. □

Thus, in particular, the replacement notion is theoretically valid for a concave program, $P_0$. However, it is important to note that concavity of problem $P_0$ is not, in general, preserved in problem $\hat{P}$. That is, the constraint set for $\hat{P}$ need not be concave. Part of the impact of this fact is reflected in the second of the following easily verified assertions for general (not necessarily concave) problems $P_0$:

i) If $\hat{x}$ is a global solution to $\hat{P}$ and $\hat{x}$ is feasible in $P_0$, then $\hat{x}$ is a global solution to $P_0$.[2,3]

1 Note that these assumptions imply neither uniqueness of $x^*$ nor convexity of $S_0$. Indeed $P_0$ could have local solutions that are not global.

2 Note that we here "assume away" the important problem of identifying I. Typical procedures for identifying I, for problems with nonlinear constraints, are ad hoc and heuristic, involving the inclusion of "blocking" constraints, the deletion of constraints with a negative Lagrange multiplier, or the utilization of penalty function maximizations to obtain points "close" to $x^*$ so that an educated estimate of I might be made. However, these procedures in general entail scaling difficulties, and they may identify a basis which is not consistent for $\hat{P}$. It would appear that the problem of identifying I is itself one of substantial difficulty.

3 Bazaraa and Goode have shown in [1] that if $\hat{x}$ is a global solution to P and if $P_0$ is a <u>strictly</u> concave program, then $\hat{x}$ will in fact be feasible in $P_0$.

ii) If $x^*$ is a local solution to $P_0$, then $x^*$ is a local solution to the corresponding[1] problem $\hat{P}$ but not conversely.[2] Specifically, suppose J is some specified subset of {1,...,m}, and suppose $\hat{x}$ is a local solution to the problem max f(x), subject to $g_i(x) = 0$, $i \in J$; then, even in the case when $P_0$ is a concave program, $\hat{x}$ need not be feasible in $P_0$; and even if $\hat{x}$ is feasible in $P_0$, it does not follow that it is a local solution to $P_0$. This holds even when J = I.

According to ii), even when $P_0$ is a concave program, one may obtain local solutions to $\hat{P}$ which are not local solutions to $P_0$. Furthermore in the case of a concave program $P_0$, when replacement by $\hat{P}$ is valid, typical penalty functions for equality constrained problems, as presented in [1] and [3], are not concave. In general it would seem obviously desirable to employ solution techniques which preserve and exploit as much of the useful structure of the original problem as possible. If f is concave and the $g_i$'s are convex then we can replace $P_0$ by either of the following problems.

| $\bar{P}$ : | $\hat{P}$ : |
|---|---|
| max f(x), subject to | max f(x), subject to |
| $g_i(x) \le 0$, $i \in I$ | $g_i(x) = 0$, $i \in I$ |

The replacement is valid in the sense that $x^*$, a global solution to $P_0$, is also a global solution to $\bar{P}$ and $\hat{P}$. If we employ a typical penalty function for equality constrained problems in solving $\hat{P}$, e.g.

$$P(x,\delta) = f(x) - \delta \sum_{i \in I} (g_i(x))^2$$

---

[1] We employ the term, corresponding, in order to emphasize the fact that problem $\hat{P}$ depends on $x^*$.

[2] Since this assertion applies to each local solution to $P_0$, it might, as suggested in a private communication by Professor Luenberger, be used to derive necessary differential optimality conditions for $P_0$. That is, such conditions are based only on local properties.

$P(x,\delta)$ is not concave in x. Thus we have lost the nice structure of problem $P_0$. Given a parameter sequence $\{\delta_n\}$, $\delta \to +\infty$, a sequence of <u>global</u> maximizers of P will converge to x* under fairly weak additional restrictions. However, the non-concavity of $\hat{P}$ raises the difficult practical problem of <u>local</u> maxima of P. All penalty functions we are aware of that are <u>specific</u> for equality constrained problems (i.e. they work for equality constraints but not inequality constraints) have this property that they do not preserve concavity. By comparison, we can choose penalty functions for $\bar{P}$ which retain the very valuable concave structure of that problem. Thus, particularly in the concave case, it seems reasonable that if the original constraints were formulated as inequalities, one might wish to use a penalty function appropriate for inequality constraints (that is, if penalty functions are to be used at all).

Finally, along a slightly different vein, it has been observed by Professor Mangasarian in a private communication that the following problem is always equivalent to $P_0$ (regardless of convexity).

$\hat{P}$ max f(x), subject to

$$g_i(x) = 0,\ i \in I$$
$$g_i(x) < 0,\ i \notin I$$

and that in reality it is this problem which one tends to solve with equality constraint procedures, taking special precaution to satisfy $g_i(x) < 0$, $i \notin I$. Typically this would be done as follows. Suppose $x_k$ is a current iterate (trial solution) for $\hat{P}$ and $I_k$ is the current guess at I. In determining the next estimate, $x_{k+1}$, suppose, for some i such that $g_i(x_k) < 0$, the algorithm is tending to force $g_i$ positive e.g. a line search is crossing the contour $g_i(x) = 0$. Then feasibility would be maintained by adding this "blocking constraint" to the basis $I_k$. A rule such as this, introduced to maintain feasibility, might retard convergence, particularly if it leads to a zig-zag phenomenon in which one constraint is repeatedly added and deleted from the basis. More explicitly, considering this same point, suppose $P_0$ is a concave program and one sets out to solve $\hat{P}$ with an equality constraint technique. Even if feasibility is maintained, as just discussed, assertion (ii) above indicates the possibility of converging to a point (a local solution to $\hat{P}$) which is not a solution to $P_0$. Provided $P_0$ is a concave program, such a false optimum will be indicated by a negative Lagrange multiplier. As first suggested by Rosen [5] in his Gradient Projection method, the algorithm can be continued by deleting any constraint with a negative multiplier, thereby effecting a change of basis. However, the nonconcavity of $\hat{P}$, resulting in local solutions which are not solutions to $P_0$, may in general lead to a waste of time in converging to wrong local optima of $\hat{P}$. Finally, if $P_0$ is not concave, then there may be Kuhn Tucker points which are not solutions, but the problem $\hat{P}$ should be expected to have more such points than $P_0$. In such cases, all one can hope for is convergence to a Kuhn Tucker point, and hence, in the nonconcave case, it should be more hazardous to try to solve $\hat{P}$.

In summary, then, even if $P_0$ is concave, $\hat{P}$ is not concave. Hence, one may have difficulties in determining global solutions to $\hat{P}$. If the initial starting point is close enough to $x^*$, and if feasibility is maintained, with anti-zig-zagging techniques, perhaps the difficulty in convergence to $x^*$ will be minimal. However, if the starting point is more remote from $x^*$, theory would expect a convergence to local solutions of $\hat{\hat{P}}$ which are not a solution to $x^*$. If $P_0$ is concave, false optima can be detected, so the issue is one of efficiency. If $P_0$ is not concave, the issue is one of reliability.

## REFERENCES


[1] Bazaraa, M.S. and J.J. Goode, "A Combinatorial Procedure for Solving Convex Programs," Working Paper, Department of Industrial Engineering, Georgia Institute of Technology,

[2] R. Fletcher, "A Class of Methods for Nonlinear Programming with Termination and Convergence Properties," Integer and Nonlinear Programming, J. Abadie, ed., North Holland, Amsterdam, 1970, pp. 161, 158.

[3] S. Lill, "A Class of Methods for Nonlinear Programming, Computational Experience with Inequality Constraints," Conference on Numerical Methods for Nonlinear Optimization, University of Dundee, June 28th-July 1, 1971, proceedings to appear.

[4] M. J. D. Powell, "A Method for Nonlinear Constraints in Minimization Problems," Optimization, R. Fletcher, ed., Academic Press, New York, 1969, p. 297.

[5] J.B. Rosen, "The Gradient Projection Method for Nonlinear Programming," Parts 1 and 2, SIAM Journal of Applied Mathematics, Vol. 8, pp. 181-217 (1960), Vol. 9, pp. 514-553 (1961).