Revised: February, 1986
(To appear in:
Discrete Applied
Mathematics)

PENALTY FORMULATION FOR ZERO-ONE NONLINEAR PROGRAMMING

B. Kalantari and J.B. Rosen

DCS-TR-148

Department of Computer Science
Busch Campus, Rutgers University
New Brunswick, NJ 08903

This research was supported in part by the National Science Foundation under Grant No. MCS-8101214 and in part by the Department of Computer Science at Rutgers University.

Penalty Formulation for Zero-One Nonlinear Programming[1]

B. KALANTARI

Department of Computer Science, Rutgers University, New Brunswick, NJ 08903, U.S.A.

J. B. ROSEN

Computer Science Department, University of Minnesota, Minneapolis, MN 55455, U.S.A.

Raghavachari has shown the equivalence of zero-one integer programming and a concave quadratic penalty function for a sufficiently large value of the penalty. A lower bound for this penalty was found by Kalantari and Rosen. It was also shown that this penalty could not be reduced in specific cases. We show that the results generalize to the case where the objective function is any concave function. Equivalent penalty formulation for non-concave functions is also considered.

Keywords: Penalty formulations, Concave minimization, Global optimization.

1. Introduction

The zero-one integer programming problem is formulated as:

$$\min_{x \in \Omega} f(x),$$

where f is linear and $\Omega = \{x: Ax \le b,\ x_i = 0,1, \text{ for } i=1,\dots,n\}$, $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and A an m x n matrix. Raghavachari [7] (see also [2]) shows the equivalence of the above problem and the following penalty formulation:

$$\min_{x \in \bar{\Omega}} w(x),$$

where $w(x) = f(x) - \mu x^T(e - x)$, with μ an appropriately chosen, sufficiently large number, e the n-vector with all elements equal to one, and $\bar{\Omega} = \{x: Ax \le b,\ 0 \le x_i \le 1, \text{ for } i=1,\dots,n\}$. Since w(x) is concave, its global minimum is attained at a vertex of $\bar{\Omega}$. If $x \in \Omega$, then $w(x) = f(x)$; furthermore if x is not an integral vertex, a large penalty is incurred by the term $\mu x^T(e - x)$, hence the equivalence of the above problems. A lower bound for μ was found in [5]. It was also shown that there exist specific cases where the lower bound cannot be improved.

---

[1] This research was supported in part by the National Science Foundation under Grant No.MCS-8101214 and in part by the Department of Computer Science at Rutgers University.

In section 2, we show that the above results hold for the more general case where $f(x)$ is an arbitrary concave function. In section 3, we consider zero-one minimization of an arbitrary function with continuous second mixed partials, and show that the problem is equivalent to a concave minimization problem.

2. Equivalent Penalty Formulation

We assume that $f(x)$ is an arbitrary concave function and consider the zero-one programming problem :

(F1): $$\min_{x \in \Omega} f(x) ,$$

where Ω is the same as before. Let $w(x) = f(x) - \mu g(x)$, where $g(x) = \sum_{i=1}^{n} (x_i - \frac{1}{2})^2$ and define

(F2): $$\min_{x \in \bar{\Omega}} w(x) ,$$

with $\bar{\Omega}$ as defined previously. Note that $w(x)$ is concave. For $x \in \Omega$, $g(x) = n/4$. It follows that when μ is sufficiently large, (F1) and (F2) are equivalent. We now obtain a lower bound for μ.

Let $\hat{\Omega}$ be the set of nonintegral vertices of $\bar{\Omega}$ and assume that $\hat{\Omega} \neq \phi$. Define

$$g_0 = \max_{x \in \hat{\Omega}} g(x),$$

$$\bar{f} = \max_{x \in \bar{\Omega}} f(x) = f(x_M),$$

$$\underline{f} = \min_{x \in \bar{\Omega}} f(x) = f(x_m), \text{ and}$$

$$\mu_0 = (\bar{f} - \underline{f})/(n/4 - g_0).$$

In (F2) pick $\mu > \mu_0$. We have

Theorem.1 Let $\tilde{x}$ be an optimal solution of (F2); then

(i) $\tilde{x} \in \Omega$ (hence an optimal solution to (F1)).
(ii) for (i) to hold, there exist specific cases where μ_0 cannot be improved.

proof. Let $x \in \Omega$ and assume $\tilde{x} \in \hat{\Omega}$. Since $g(\tilde{x}) \leq g_0$, we have

$$w(x) - w(\tilde{x}) \leq f(x) - f(\tilde{x}) - \mu(n/4 - g_0) .$$

Since $\mu > \mu_0$ and $(n/4 - g_0) \geq 0$, we get

$w(x) - w(\tilde{x}) < f(x) - f(\tilde{x}) - (\bar{f} - \underline{f}) = (f(x) - \bar{f}) - (f(\tilde{x}) - \underline{f}) \leq 0.$
This contradicts the optimality of $\tilde{x}$ and hence
$\tilde{x} \in \Omega$. The proof of (i) is then complete by noting that
$f(\tilde{x}) - f(x) = w(\tilde{x}) - w(x) \leq 0.$

To prove (ii), consider the case where x_M is the only integral vertex of $\bar{\Omega}$, $g_0 = \underline{f} = g(x_m)$, and $w(x) = f(x) - \mu' g(x)$ with $0 \leq \mu' < \mu_0$; then

$w(x_M) - w(x_m) = \bar{f} - \underline{f} - \mu'(n/4 - g_0) > 0$. Hence the proof of (ii). ■

3. Penalty Formulation for Non-concave Functions

In this section we consider the following zero-one programming problem:

(G1): $\quad \min_{x \in \Omega} f(x),$

where f is not necessarily concave. We assume that the second mixed partials of f(x) are continuous over $\bar{\Omega}$. We show that problem (G1) is equivalent to a zero-one concave minimization. For x in $\bar{\Omega}$ let $\mu_M(x)$ be the maximum of the absolute values of the eigenvalue of H(x), the Hessian of f(x). It is well-known that for any operator matrix norm $\|\cdot\|$, we have

$$\mu_M(x) \leq \| H(x) \|$$

(see for example [1, p. 412]). The above inequality, the compactness of $\bar{\Omega}$ together with the continuity assumption imply that $\mu_M = \sup_{x \in \Omega} \mu_M(x) < \infty$. Define $F(x) = f(x) - 1/2\mu_M g(x)$, where g(x) as before is given by $g(x) = \sum_{i=1}^{n} (x_i - \frac{1}{2})^2$. The Hessian of F(x) is given by $H'(x) = H(x) - \mu_M I$, where I is the n by n identity matrix. H'(x) is negative semidefinite, hence F(x) is concave and problem (G1) is equivalent to the following zero-one programming:

(G2): $\quad \min_{x \in \Omega} F(x).$

Now since F(x) is concave, (G2) may be replaced by an equivalent continuous relaxation as described in the previous section.

Remark. From a practical point of view, for an arbitrary function f(x) and an arbitrary Ω, the calculation of the quantities $\bar{f}$, $\underline{f}$, g_0, and μ_M may require as much time as that required to solve the original problem. However, in specific cases the concave formulation may be obtained with a reasonable amount of calculation. For example a problem of considerable interest is the special case where f(x) is a quadratic function and $\bar{\Omega}$ the entire hypercube (see for example [3] and the survey paper [4]). In this case, one only needs to compute μ_M and this can be done easily. Now once such a concave formulation is at hand, one may employ methods developed for the global minimization of concave quadratic functions. For this special case, a branch and bound method is described in [6] where the lower bounding scheme is based on such concave formulation.

4. References

1. K. E. Atkinson, An Introduction to Numerical Analysis (John Wiley and Sons, New York 1978).
2. A. Charnes and W. W. Cooper, Management Models and Industrial Applications of Linear Programming (John Wiley and Sons, New York 1961).
3. P. L. Hammer, P. Hansen, B. Simeone, Roof duality, complementation and persistency in quadratic 0-1 optimization, Math. Prog. 28(1984) 121-155.
4. P. Hansen, Methods for nonlinear 0-1 programming, Ann. Discrete Math. 5(1979), 53-70.
5. B. Kalantari and J. B. Rosen, Penalty for zero-one integer equivalent problem, Math. Prog. 24(1982) 229-232.
6. B. Kalantari, Large scale global minimization of linearly constrained concave quadratic functions and Related problems, Ph.D. thesis, Computer Science Department, University of Minnesota (1984).
7. M. Raghavachari, On connection between zero-one integer programming and concave minimization under linear constraints, Operations Research 17(1969) 680- 684.