Penalty Methods in Constrained Optimization

Parwadi Moengin

***Abstract*—In this work, we study a class of polynomial order-even penalty functions for solving equality constrained optimization problem with the essential property that each member is convex polynomial order-even when viewed as a function of the multiplier. Under certain assumption on the parameters of the penalty function, we give a rule for choosing the parameters of the penalty function. We also give an algorithm for solving this problem.**

***Index Terms*—constrained optimization, penalty method, polynomial order-even.**

I. INTRODUCTION

The basic idea in penalty method is to eliminate some or all of the constraints and add to the objective function a penalty term which prescribes a high cost to infeasible points.

Associated with this method is a parameter σ, which determines the severity of the penalty and as a consequence the extent to which the resulting unconstrained problem approximates the original constrained problem. In this paper, we restrict attention to the polynomial order-even penalty function. Other penalty functions will appear elsewhere.

II. STATEMENT OF THE PROBLEM

Throughout this paper we consider the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to } \ f_i(x) = 0, \text{ for } I = 1,\ldots,m \ \ x \in X, \end{aligned} \tag{1}$$

where f, $f_1,\ldots,f_m$ be real-valued function on a set $X \subset R^n$. We assume that problem (1) has at least one feasible solution.

III. POLYNOMIAL ORDER-EVEN PENALTY FUNCTION

For any scalar σ > 0, let us define the polynomial order-even penalty function

$$P(x,\sigma): R^n \to R$$

by

$$P(x,\sigma) = f(x) + \sigma \sum_{i=1}^{m} (f_i(x))^\rho , \tag{2}$$

where ρ > 0 is even number. The positive even number ρ is chosen to ensure that the function (2) is convex. Hence, $P(x,\sigma)$ has a global minimum. We refer to σ as the penalty parameter. In fact, this is just the ordinary Lagrange function for the altered problem in which the constraints $f_1,\ldots,f_m$ are replaced by $f_1^\rho,\ldots,f_m^\rho$.

The motivation behind the introduction of the polynomial order-term that they may lead to a representation of a local optimal solution in terms of a local unconstrained minimum.

The polynomial order-even penalty method consists of solving a sequence of problems of the form

$$\begin{aligned} &\text{minimize} \quad P(x,\sigma^k) \\ &\text{subject to } \ x \in X, \end{aligned} \tag{3}$$

where $\{\sigma^k\}$ is a penalty parameter sequence satisfying

$$0 < \sigma^k < \sigma^{k+1} \text{ for all } k, \ \ \sigma^k \to \infty .$$

The method depends for its success on sequentially increasing the penalty parameter to infinity. In this paper, we concentrate on the effect of the penalty parameter.

The rationale for the penalty method is based on the fact that when $\sigma^k \to \infty$, then the term

$$\sigma^k \sum_{i=1}^{m} (f_i(x))^\rho ,$$

which is added to the objective function, tends to infinity if $f_i(x) \neq 0$ and equals zero if $f_i(x) = 0$. Thus, we define the function

$$\hat{f} : R^n \to (-\infty, +\infty]$$

by

$$\hat{f}(x) = \begin{cases} f(x) & \text{if } \ f_i(x) = 0 \ \text{ for all } i, \\ \infty & \text{if } \ f_i(x) \neq 0 \ \text{ for all } i, \end{cases}$$

the optimal value of the original problem can be written as

$$\begin{aligned} f^* &= \inf_{\substack{f_i(x)=0 \\ x \in X}} f(x) = \inf_{x \in X} \hat{f}(x) \\ &= \inf_{x \in X} \lim_{k \to \infty} P(x,\sigma^k). \end{aligned} \tag{4}$$

On the other hand, the penalty method determines, via the sequence of minimizations (3),

$$\bar{f} = \lim_{k \to \infty} \inf_{x \in X} P(x,\sigma^k). \tag{5}$$

Thus, in order for the penalty method to be successful, the original problem should be such that the interchange of "lim" and "inf" in (4) and (5) is valid.

The following theorem guarantees the validity of the interchange, under mild assumptions, and constitutes the basic convergence result for the penalty method.

Manuscript received September 20, 2007. This work was supported by the Faculty of Industrial Engineering, Trisakti University, Jakarta.

Parwadi Moengin is with the Department of Industrial Engineering, Faculty of Industrial Technology, Trisakti University, Jakarta 11440, Indonesia. (e-mail: parwadi@ hotmail.com, parwadi@trisakti.ac.id).

First, we derive the convexity behavior of the polynomial penalty function defined by (2) is stated in the following stated theorem.

Theorem 1 (Convexity)

The polynomial penalty function $P(x,\sigma)$ *is convex in its domain for every* $\sigma > 0$.

Proof.
It is straightforward to prove convexity of $P(x,\sigma)$ using the convexity of $c^T x$ and $(f_i(x))^\rho$. Then the theorem is proven. ■

The local and global behavior of the polynomial penalty function defined by (3.1) is stated in next the theorem. It is a consequence of Theorem 1.

Theorem 2 (Local and global behavior)

Consider the function $P(x,\sigma)$ *which is defined in (3.1). Then*

(a) $P(x,\sigma)$ *has a finite unconstrained minimizer in its domain for every* $\sigma > 0$ *and the set* M_σ *of unconstrained minimizers of* $P(x,\sigma)$ *in its domain is convex and compact for every* $\sigma > 0$.

(b) Any unconstrained local minimizer of $P(x,\sigma)$ *in its domain is also a global unconstrained minimizer of* $P(x,\sigma)$.

Proof.

It follows from Theorem 1 that the smooth function $P(x,\sigma)$ achieves its minimum in its domain. We then conclude that $P(x,\sigma)$ has at least one finite unconstrained minimizer.

By Theorem 1 $P(x,\sigma)$ is convex, so any local minimizer is also a global minimizer. Thus, the set M_σ of unconstrained minimizers of $P(x,\sigma)$ is bounded and closed, because the minimum value of $P(x,\sigma)$ is unique, and it follows that M_σ is compact. Clearly, the convexity of M_σ follows from the fact that the set of

optimal points $P(x,\sigma)$ is convex. Theorem 2 has been verified. ■

As a consequence of Theorem 2, we derive the monotonicity behaviors of the objective function problem (P), the penalty terms in $P(x,\sigma)$ and the minimum value of the primal polynomial penalty function $P(x,\sigma)$.

To do this, for any $\sigma^k > 0$ we denote x^k and $P(x^k,\sigma^k)$ as a minimizer and minimum value of problem (2), respectively.

Theorem 3
Assume that f and f_i *are continuous functions and X a closed set. For k = 0, 1, …, let* $\{x^k\}$ *be a global minimum of the problem* (3), *where* $0 < \sigma^k < \sigma^{k+1}$ *for all k,* $\sigma^k \to \infty$. *Then every limit point of the sequence* (x^k) *is a global minimum of the problem* (1).

Proof.
We have by definition of x^k

$$P(x^k,\sigma^k) \le P(x,\sigma^k) \text{ for all } x \in X. \quad (6)$$

Let f^* denote the optimal value of the original problem. We have

$$f^* = \inf_{\substack{f_i(x)=0\\ x\in X}} f(x) = \inf_{\substack{f_i(x)=0\\ x\in X}} P(x,\sigma^k).$$

Hence, by taking the infimum of the right-hand side of (6) over $x \in X$, $f_i(x) = 0$, for $i = 1,\dots,m$, we obtain

$$P(x^k,\sigma^k) = f(x^k) + \sigma^k \sum_{i=1}^{m} (f_i(x^k))^\rho \le f^*.$$

Let $\bar{x}$ be a limit point of $\{x^k\}$. By taking the limit superior in the above relation and by using the continuity of f and f_i, we obtain

$$f(\bar{x}) + \limsup_{k\to\infty} \sigma^k \sum_{i=1}^{m} (f_i(x^k))^\rho \le f^*. \quad (7)$$

Since

$$\sum_{i=1}^{m} \left(f_i(x^k)\right)^\rho \ge 0,\ \sigma^k \to \infty,$$

it follows that we must have

$$\sum_{i=1}^{m} (f_i(x^k))^\rho \to 0$$

and

$$f_i(\bar{x}) = 0 \text{ for all } i = 1,\dots,m, \quad (8)$$

for otherwise the limit superior in the left-hand side of (7) will equal $+\infty$. Since X is a closed set we also obtain that $\bar{x} \in X$. Hence, $\bar{x}$ is feasible, and

$$f^* \le f(\bar{x}). \quad (9)$$

Using (7)-(9), we obtain

$$f^* + \limsup_{k\to\infty} \sigma^k \sum_{i=1}^{m} (f_i(x^k))^\rho \le$$

$$f(\bar{x}) + \limsup_{k\to\infty} \sigma^k \sum_{i=1}^{m} (f_i(x^k))^\rho \le f^*.$$

Hence,

$$\limsup_{k\to\infty} \sigma^k \sum_{i=1}^{m} (f_i(x^k))^\rho = 0$$

and

$$f(\bar{x}) = f^*,$$

which proves that $\bar{x}$ is a global minimum for problem (1). ■

IV. ALGORITHM

The implications of these theorems are remarkably strong. The polynomial penalty method has a finite unconstrained minimizer for every value of the penalty parameter, and every limit point of a minimizing sequence for the penalty function is a constrained minimizer of problem (1). Thus the algorithm of solving a sequence of minimization problems is suggested.

1. Given σ^k to get x^k by minimizing problem (3). (Note: Some techniques are available to determine a minimum x^k of $P(x,\sigma^k)$, such as calculus technique, Newton-Raphson numerical approach, and so on.).

2. Then set σ^{k+1}, where $\sigma^{k+1} > \sigma^k$, and $\sigma^k \to \infty$ to get a sequence $\{x^k\}$.
3. According to Theorem 3 every limit point of the sequence is a global minimum of problem (1).

Example

Consider the following problem

minimize x
subject to $x = 1$.

Note that the minimum value of the problem is 1 at the minimum point $x^* = 1$. According to Eq. (2), we have a sequence of polynomial order-2 penalty function (hence, $\rho = 2$)

$$P(x,\sigma^k) = x + \sigma^k (x-1)^2,\ k = 1,2,\ldots,$$

and $\sigma > 1$. The derivative of $P(x,\sigma^k)$ with respect to x is given by

$$P'(x,\sigma^k) = 2\sigma^k x + \left(1 - 2\sigma^k\right).$$

The minimum point of $P(x,\sigma^k)$ is obtained by solving the equation $P'(x,\sigma^k) = 0$. Therefore, the minimum point of $P(x,\sigma^k)$ is given by

$$x^k = \frac{2\sigma^k - 1}{2\sigma^k} = 1 - \frac{1}{2\sigma^k},\ k = 1,2,\ldots\ .$$

If $\sigma = 2$, we have a sequence

$$\{x^k\} = \left\{1 - \frac{1}{2^{k+1}}, k = 1,2,3,\ldots\right\},$$

and the limit point of the sequence is $\bar{x} = x^* = 1$ as $k \to \infty$. With the same way, if $\sigma = 5$, we have another sequence

$$\{x^k\} = \left\{1 - \frac{1}{5(2^k)}, k = 1,2,3,\ldots\right\},$$

and the limit point of the new sequence is also $\bar{x} = x^* = 1$ as $k \to \infty$. ■

However, Theorem 3 shown above has several weaknesses. First, it assumes that the problem (3) has a global minimum. This may not true, even if the original problem (1) has a global minimum. As an example, consider the scalar problem

minimize $-x^6$
subject to $x = 0$.

This problem has a unique global minimum at the point $x^* = 0$. When we choose $\rho = 2$, we have

$$P(x,\sigma^k) = -x^6 + \sigma^k x^2.$$

Clearly, $\inf_x P(x,\sigma^k) = -\infty$, and hence $P(x,\sigma^k)$ has no global minimum for every σ^k. The same result is obtained when we choose $\rho = 4$. This example shows a weakness of the penalty method. However, $P(x,\sigma^k)$ has global minimum for every σ^k when we take $\rho \ge 6$.

Note that the closed assumption to the set X is important to ensure that the limit point of the sequence $\{x^k\}$ to be in X.

We also note that if we add boundedness assumption to the closed set X (hence, X is a compact set) in the theorem, then $P(x,\sigma^k)$ will attain a global minimum over X. If it occurs, the problem (3) must has a global minimum.

For the inequality-constrained problem

minimize $f(x)$
subject to $f_i(x) \le 0$, for $i = 1,\ldots,m$

$$x \in X, \tag{10}$$

it is not immediately apparent what form the penalty function should have. The study of this problem will appear elsewhere.

REFERENCES

[1] Durazzi, C. (2000. On the Newton interior-point method for nonlinear programming problems. *Journal of Optimization Theory and Applications*. 104(1). pp. 73−90.

[2] Kas, P., Klafszky, E., & Malyusz, L. (1999). Convex program based on the Young inequality and its relation to linear programming. *Central European Journal for Operations Research*. 7(3). pp. 291−304.

[3] Parwadi, M. (2002). *The Exponential Function Methods for Solving LP Problems*. Paper presented at the Seminar in Department of Mathematics, Universiti Putra Malaysia.

[4] Parwadi, M., Mohd, I.B., & Ibrahim, N.A. (2002). *Solving Bounded LP Problems using Modified Logarithmic-exponential Functions*. In Purwanto (Ed.), Proceedings of the National Conference on Mathematics and Its Applications in UM Malang (pp. 135-141). Malang: Department of Mathematics UM Malang.

[5] Singiresu, S.R. (1996). *Engineering Optimization Theory and Practice* (3rd ed.). New York: Wiley & Sons.

[6] Wright, S.J. (2001). *On the convergence of the Newton/log-barrier method*. Mathematical Programming, 90(1), 71−100.

[7] Zboo, R.A, Yadav, S.P., & Mohan, C. (1999). Penalty method for an optimal control problem with equality and inequality constraints. *Indian Journal of Pure and Applied Mathematics*. 30(1), pp. 1−14.