Polynomial Penalty Method for Solving Linear Programming Problems

Parwadi Moengin, *Member, IAENG*

***Abstract*—In this work, we study a class of polynomial order-even penalty functions for solving linear programming problems 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*—linear programming, 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 (Wright, 2001; Zboo, etc., 1999). 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 problem (Kas, etc., 1999; Parwadi, etc., 2002). In this paper, we restrict attention to the polynomial order-even penalty function. Other penalty functions will appear elsewhere. This paper is concerned with the study of the polynomial penalty function methods for solving linear programming. It presents some background of the methods for the problem. The paper also describes the theorems and algorithms for the methods. At the end of the paper we give some conclusions and comments to the methods.

II. STATEMENT OF THE PROBLEM

Throughout this paper we consider the problem

$$\begin{aligned} &\text{minimize} \quad c^T x \\ &\text{subject to } Ax = b \\ &\qquad x \geq 0, \end{aligned} \qquad (1)$$

where $A \in R^{m\times n}$, $c, x \in R^n$, and $b \in R^m$. Without loss of generality we assume that *A* has full rank *m*. We assume that problem (1) has at least one feasible solution. In order to solve this problem, we can use Karmarkar's algorithm and simpelx method (Durazzi, 2000). But in this paper we propose a polynomial penalty method as another alternative method to solve linear programming problem (1).

Manuscript is submitted on March 6, 2010. 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. (email: parwadi@hotmail.com, parwadi@trisakti.ac.id).

III. POLYNOMIAL PENALTY METHOD

For any scalar σ > 0, we define the polynomial penalty function $P(x,\sigma)$ for problem (1); $P(x,\sigma)$: $R^n \to R$ by

$$P(x,\sigma) = c^T x + \sigma \sum_{i=1}^{m} (A_i x - b_i)^\rho, \qquad (2)$$

where ρ > 0 is an even number. Here, A_i and b_i denote the *i*th row of matrices *A* and *b*, respectively. 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.

This is the ordinary Lagrangian function in which in the altered problem, the constraints $A_i x - b_i$ (*i* =1,..,*m*) are replaced by $(A_i x - b_i)^\rho$. The penalty terms are formed from a sum of polynomial order–ρ of constrained violations and the penalty parameter σ determines the amount of the penalty.

The motivation behind the introduction of the polynomial order-ρ term is that they may lead to a representation of an optimal solution of problem (1) in terms of a local unconstrained minimum. Simply stating the definition (2) does not give an adequate impression of the dramatic effects of the imposed penalty. In order to understand the function stated by (2) we give an example with some values for σ. Some graphs of $P(x,\sigma)$ are given in Figures 1–3 for the trivial problem

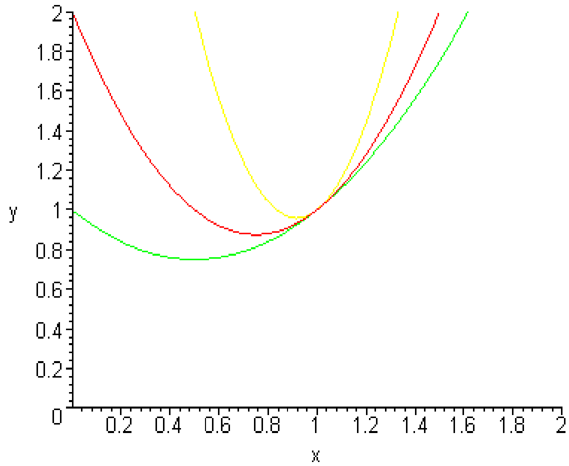
$$\begin{aligned} &\text{minimize } f(x) = x \\ &\text{subject to } x - 1 = 0, \end{aligned}$$

for which the polynomial penalty function is given by

$$P(x,\sigma) = x + \sigma(x-1)^\rho .$$

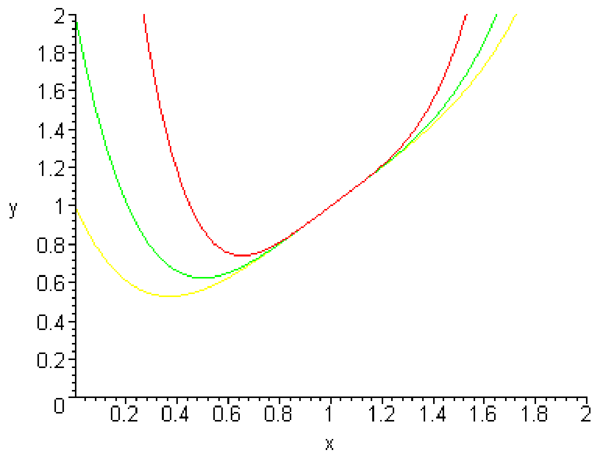
Figures 1, 2 and 3 depict the one-dimensional variation of the penalty function of ρ = 2, 4 and 6, for three values of penalty parameter σ, that is σ = 1, σ = 2 and σ = 6, respectively.

(Advance online publication: 19 August 2010)



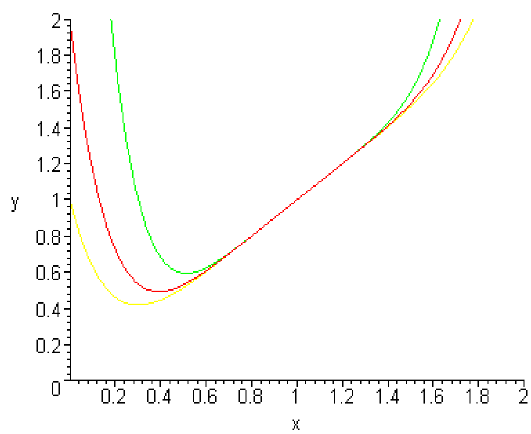
x(1) *x*(2) *x*(6)

Figure 1 The quadratic penalty function for ρ = 2



x(1) *x*(2) *x*(6)

Figure 2 The polynomial penalty function for ρ = 4



x(1) *x*(2) *x*(6)

Figure 3 The polynomial penalty function for ρ = 6

The *y*-ordinates of these figures represent $P(x,\sigma)$ for ρ = 2, ρ = 4, ρ = 6, respectively. Clearly, if the solution x^* = 1 of this example is compared with the points which minimize $P(x,\sigma)$, it is clear that x^* is a limit point of the unconstrained minimizers of $P(x,\sigma)$ as $\sigma \to \infty$.

The intuitive motivation for the penalty method is that we seek unconstrained minimizers of $P(x,\sigma)$ for value of σ increasing to infinity. Thus the method of solving a sequence of minimization problem can be considered.

The polynomial penalty method for problem (1) consists of solving a sequence of problems of the form

$$\begin{aligned} &\text{minimize } P(x,\sigma^k) \\ &\text{subject to } \; x \geq 0, \end{aligned} \tag{3}$$

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

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

when added to the objective function, tends to infinity if $A_i x - b_i \neq 0$ and equals zero if $A_i x - b_i = 0$ for all *i*. Thus, we define the function $f : R^n \to (-\infty, +\infty]$ by

$$f(x) = \begin{cases} c^T x & \text{if } \; A_i x - b_i = 0 \text{ for all } i, \\ \infty & \text{if } \; A_i x - b_i \neq 0 \text{ for all } i. \end{cases}$$

The optimal value of the original problem (1) can be written as

$$\begin{aligned} f^* &= \inf_{Ax=b, x\geq 0} c^T x = \inf_{x \geq 0} f(x) \\ &= \inf_{x\geq 0} \lim_{k\to\infty} P(x,\sigma^k). \end{aligned} \tag{4}$$

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

$$\bar{f} = \lim_{k\to\infty} \inf_{x\geq 0} P(x,\sigma^k). \tag{5}$$

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. Before we give a guarantee for the validity of the interchange, we investigate some properties of the function defined in (2).

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 $(A_i x - b_i)^\rho$. Then the theorem is proven. ■

The local and global behavior of the polynomial penalty function defined by (2) 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 (2). 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 (1), the penalty terms in $P(x,\sigma)$ and the minimum value of the 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 (3), respectively.

Theorem 3 (Monotonicity)
Let $\{\sigma^k\}$ be an increasing sequence of positive penalty parameters such that $\sigma^k \to \infty$ as $k \to \infty$.
Then

(a) $\{c^T x^k\}$ is non-decreasing.

(b) $\left\{\sum_{i=1}^{m} (A_i x^k - b_i)^\rho\right\}$ is non-increasing.

(c) $\{P(x^k,\sigma^k)\}$ is non-decreasing.

Proof.
Let x^k and x^{k+1} denote the global minimizers of problem (3) for the penalty parameters σ^k and σ^{k+1}, respectively. By definition of x^k and x^{k+1} as minimizers and $\sigma^k < \sigma^{k+1}$, we have

$$c^T x^k + \sigma^k \sum_{i=1}^{m} (A_i x^k - b_i)^\rho \le c^T x^{k+1} + \sigma^k \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho , \qquad (6a)$$

$$c^T x^{k+1} + \sigma^k \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho \le c^T x^{k+1} + \sigma^{k+1} \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho , \qquad (6b)$$

$$c^T x^{k+1} + \sigma^{k+1} \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho \le c^T x^k + \sigma^{k+1} \sum_{i=1}^{m} (A_i x^k - b_i)^\rho . \qquad (6c)$$

We multiply the first inequality (6a) with the ratio σ^{k+1}/σ^k, and add the inequality to the inequality (6c) we obtain

$$\left(\frac{\sigma^{k+1}}{\sigma^k} - 1\right) c^T x^k \le \left(\frac{\sigma^{k+1}}{\sigma^k} - 1\right) c^T x^{k+1} .$$

Since $0 < \sigma^k < \sigma^{k+1}$, it follows that $c^T x^k \le c^T x^{k+1}$ and part (a) is established.

To prove part (b) of the theorem, we add the inequality (6a) to the inequality (6c) to get

$$\left(\sigma^{k+1} - \sigma^k\right) \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho \le \left(\sigma^{k+1} - \sigma^k\right) \sum_{i=1}^{m} (A_i x^k - b_i)^\rho ,$$

thus

$$\sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho \le \sum_{i=1}^{m} (A_i x^k - b_i)^\rho$$

as required for part (b).

Using inequalities (6a) and (6b), we obtain

$$c^T x^k + \sigma^k \sum_{i=1}^{m} (A_i x^k - b_i)^\rho \le c^T x^{k+1} + \sigma^{k+1} \sum_{i=1}^{m} (A_i x^{k+1} - b_i)^\rho .$$

Hence, part (c) of the theorem is established. ■

We now give the main theorem concerning polynomial penalty method for linear programming problem (1).

Theorem 4 (Convergence of polynomial penalty function)

Let $\{\sigma^k\}$ be an increasing sequence of positive penalty parameters such that $\sigma^k \to \infty$ as $k \to \infty$. Denote x^k and $P(x^k, \sigma^k)$ as in Theorem 3. Then

(a) $Ax^k \to b$ as $k \to \infty$.

(b) $c^T x^k \to f^*$ as $k \to \infty$.

(c) $P(x^k, \sigma^k) \to f^*$ as $k \to \infty$.

Proof.

By definition of x^k and $P(x^k, \sigma^k)$, we have

$c^T x^k \le P(x^k, \sigma^k) \le P(x, \sigma^k)$ for all $x \ge 0$. (7)

Let f^* denotes the optimal value of the problem (P). We have

$$f^* = \inf_{Ax=b, x\ge 0} c^T x = \inf_{\substack{Ax=b\\ x\ge 0}} P(x, \sigma^k).$$

Hence, by taking the infimum of the right-hand side of (7) over $x \ge 0$ and $Ax = b$, we obtain

$$P(x^k, \sigma^k) = c^T x^k + \sigma^k \sum_{i=1}^{m} (A_i x^k - b_i)^\rho \le f^*.$$

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 $c^T x$ and $A_i x - b_i$, we obtain

$$c^T \bar{x} + \limsup_{k\to\infty} \sigma^k \sum_{i=1}^{m} (A_i x^k - b_i)^\rho \le f^*. \quad (8)$$

Since $\sum_{i=1}^{m} (A_i x^k - b_i)^\rho \ge 0$ and $\sigma^k \to \infty$, it follows that we must have

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

and

$$A_i \bar{x} - b_i = 0 \text{ for all } i = 1, \ldots, m, \quad 9)$$

otherwise the limit superior in the left-hand side of (8) will equal to $+\infty$. This proves part (a) of the theorem.

Since $\{x \in R^n \mid x \ge 0\}$ is a closed set we also obtain that $\bar{x} \ge 0$. Hence, $\bar{x}$ is feasible, and

$$f^* \le c^T \bar{x}. \quad (10)$$

Using (8)-(10), we obtain

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

Hence,

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

and

$$f^* = c^T \bar{x},$$

which proves that $\bar{x}$ is a global minimum for problem (1). This proves part (b) of the theorem.

To prove part (c), we apply the results of parts (a) and (b), and then taking $k \to \infty$ of the definition $P(x^k, \sigma^k)$. ■

Some notes about this theorem will be taken. First, it assumes that the problem (3) has a global minimum. This may not be true if the objective function of the problem (1) is replaced by a nonlinear function. However, this situation may be handled by choosing appropriate value of ρ. We also note that the constraint $x \ge 0$ of the problem (3) is important to ensure that the limit point of the sequence $\{x^k\}$ satisfies the condition $x \ge 0$.

IV. ALGORITHM

The implications of these theorems are remarkably strong. The polynomial penalty function 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 a problem (1). Thus the algorithm of solving a sequence of minimization problems is suggested. Based on Theorems 4, we formulate an algorithm for solving problem (1).

Algorithm 1

Given $Ax = b$, $\sigma^1 > 0$, the number of iteration N and $\varepsilon > 0$.

1. Choose $x^1 \in R^n$ such that $Ax^1 = b$ and $x^1 \ge 0$.
2. If the optimality conditions are satisfied for problem (1) at x^1, then stop.
3. Compute $P(x^1, \sigma^1) := \min_{x\ge 0} P(x, \sigma^1)$ and the minimizer x^1.
4. Compute $P(x^k, \sigma^k) := \min_{x\ge 0} P(x, \sigma^k)$, the minimizer x^k and $\sigma^k := 10\,\sigma^{k-1}$ for $k = 2$.
5. If $\| x^k - x^{k-1} \| < \varepsilon$ or $| P(x^k, \sigma^k) - P(x^{k-1}, \sigma^{k-1}) | < \varepsilon$ or

$\| Ax^k - b \| < \varepsilon$ or $k = N$; then stop.

Else $k := k + 1$ and go to step 4. ■

Examples: Consider the following problems.

1.Minimize $f = 2x_1 + 5x_2 + 7x_3$
subject to $x_1 + 2x_2 + 3x_3 = 6$,
$x_j \geq 0$, for j = 1, 2, 3.

2.Minimize $f = 0.4x_1 + 0.5x_2$
subject to $0.3x_1 + 0.1x_2 \geq 2.7$,
$0.5x_1 + 0.5x_2 = 6$,
$x_j \geq 0$, for j = 1, 2.

3.Minimize $f = 4x_1 + 3x_2$
subject to $2x_1 + 3x_2 \geq 6$,
$4x_1 + x_2 \geq 4$,
$x_j \geq 0$, for j = 1, 2.

4.Minimize $f = -3x_1 + 4x_2$
subject to $x_1 - x_2 \geq 0$,
$-x_1 + 2x_2 \geq 2$,
$x_j \geq 0$, for j = 1, 2.

5. Minimize $f = 3x_1 + 8x_2$
subject to $3x_1 + 4x_2 \leq 20$,
$x_1 + 3x_2 \geq 12$,
$x_j \geq 0$, for j = 1, 2.

Table 1 reports the results of computational for Algorithm 1(ρ = 2), Algorithm 1(ρ = 4) and Karmarkar's Algorithm. The first column of Table 1 contains the problem number and the next two columns of each algorithm in this table contain the total iterations and the times (in seconds) of each algorithm.

Tabel 1 Algorithm 1(ρ = 2), Algorithm 1(ρ = 4) and Karmarkar's Algorithm test statistics

Problem No.	Algorithm 1(ρ = 2)		Algorithm 1(ρ = 4)		Karmarkar's Algorithm	
	Total Iterations	Time (Secs.)	Total Iterations	Time (Secs.)	Total Iterations	Time (Secs.)
1.	10	3.4	25	189.1	16	3.6
2.	9	3.9	7	75.3	19	3.7
3.	9	8.5	9	851.4	19	3.7
4.	10	8.7	12	2978.9	12	2.8
5.	10	9.9	19	5441.7	18	3.8

V. CONCLUSION

As mentioned above, the paper has described the penalty functions with penalty terms in polynomial order-σ for solving problem (1). The algorithms for these methods are also given in this paper. The Algorithm 1 is used to solve the problem (1). We also note the important thing of these methods which do not need an interior point assumption.

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., 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.

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

[5] 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.