TAIWANESE JOURNAL OF MATHEMATICS
Vol. 18, No. 3, pp. 809-815, June 2014
DOI: 10.11650/tjm.18.2014.3349
This paper is available online at http://journal.taiwanmathsoc.org.tw

ANALYSIS OF PENALTY PARAMETERS IN BINARY CONSTRAINED EXTREMUM PROBLEMS

Carla Antoni and Mario Pedrazzoli

Abstract. A given discrete constrained extremum problem can be associated with an equivalent continuous one, the equivalence being assured by the equality of their sets of solutions. When the equivalent problem is a penalized one, a crucial question is the size of the penalty parameter. The present paper concerns the case where the problem is a 0-1 extremum (linear) one.

1. Introduction and Notation

Consider the constrained extremum problem

$$\min f(x) \ \text{ s.}t. \ \ x \in K, \tag{1}$$

where f is a real-valued function on a compact set $K \subset \mathbb{R}^n$. A problem with the same set of solutions of (1) is said equivalent to (1). The replacement of (1) with an equivalent one can be profitable if, for example, (1) is a discrete problem and the equivalent problem is a continuous one.

An equivalent problem of (1) can be of the kind

$$\min f(x) \ \text{ s.}t. \ \ x \in K^*, \tag{2}$$

where K^* is a compact supset of K; in this case (2) is called a relaxed problem of (1). Sometimes (2) can be solved more easily than (1) and, if a solution of (2) belongs to K, this point solves (1) too. If this does not happen, other methods can be adopted: among them, exact penalty methods.

A penalized problem miminizes a funtion $f + \mu\Phi$ on a compact K^* including K, Φ being a suitable penalty function and μ being a suitable positive real parameter, called penalty parameter; statements of exact penalty show the penalized problem

$$\min \Big[f(x) + \mu\Phi(x) \Big] \ \text{ s.}t. \ \ x \in K^* \tag{3}$$

Received May 3, 2013, accepted November 7, 2013.
Communicated by Franco Giannessi.
2010 *Mathematics Subject Classification*: 05B99, 65K05, 90Cxx, 90C05, 90C09, 90C10.
Key words and phrases: Penalty problem, Relaxed problem, Exact penalty, Penalty parameter, Penalty function.

has the same set of solutions of the given one (1). In this approaches, a first crucial question is to avoid to work with too large μ; the best goal is to find the minimum μ_0 such that, for every μ greater than μ_0, (1) and (3) are equivalent. If this goal cannot be gained easily, it is significant to find a bound $\bar{\mu}$ such that, for every μ greater than $\bar{\mu}$, the equivalence of the problems is assured; in this case, a possible bound less than $\bar{\mu}$ represents an improvement of the previous one. To this aim, in this paper, the parameter introduced in [4], for a zero-one problem is improved starting from results of [2].

The considered problem is the zero-one problem dealt in [4] by Kalantari and Rosen:

(4) $$\min\,(-c^t x) \;\text{ s.}t.\; x \in R \cap Z,$$

where R is a polyhedron of $\mathbb{R}^n$, $Z = \{0,1\}^n$ is the set of the vertices of the set $X = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1\}$, c and x are in $\mathbb{R}^n$ and c^t denotes the transpose of c; without loss of generality, c is supposed to have non-negative components.

The approach proposed in [4] starts from the relaxed problem

(5) $$\min\,(-c^t x) \;\text{ s.}t.\; x \in R \cap X;$$

if a solution x_0 of (5) does not belong to Z, the problem

(6) $$\max\; g(x) \;\text{ s.}t.\; x \in \hat{F},$$

where $g(x) = \sum_{i=1}^n (x_i - 1/2)^2$, and $\hat{F}$ denotes the set of the vertices of $R \cap X$ not in Z, is considered. Denoted by $\bar{x}$ a solution of (6), the equivalence between (4) and

(7) $$\min\,\big[- c^t x - \mu g(x)\big] \;\text{ s.}t.\; x \in R \cap X,$$

when $\mu > \mu_R$, is proved, being

$$\mu_R = \frac{c^t x_0}{n/4 - g(\bar{x})}.$$

In order to improve a result of [4], we refer to [2], where exact penalty results concern more general cases: the problem (4) satisfies the hypothesis of Theorem 3.1 of [2], then the equivalence between (4) and the following penalized extremum problem

(8) $$\min\,\big[- c^t x + \mu\varphi(x)\big] \;\text{ s.}t.\; x \in R \cap X,$$

for a suitable function φ and parameter μ holds. More precisely, φ is the real-valued function defined by

(9) $$\varphi(x) = \sum_{i=1}^{n} x_i(1 - x_i),$$

and μ is any real parameter greater than μ_G, where μ_G is the amount that will be described. Observe that, since $\varphi = n/4 - g$, the penalty problems (7) and (8) differ only for the constant $\mu \cdot n/4$.

In the sequel, the following notation is adopted.

For a subset A of $\mathbb{R}^n$, A^c denotes the complement of A with respect to X; for a point $x \in \mathbb{R}^n$, $d(x, A)$ denotes the euclidean distance between x and the set A; for $y \in \mathbb{R}^n$ and $r \in \mathbb{R}_+$, $B(y, r) = \{x \in \mathbb{R}^n : d(x, y) < r\}$.

Let $\rho \in]0, 1[$ and $\{z_i,\ i = 1, \ldots, k\} = R \cap Z$; put $X_\rho := (X \cap R) \setminus \Big(\cup_{i=1}^k B(z_i, \rho)\Big)$, and, finally, if

$$\lambda_\rho := \frac{\max_{x \in R \cap X} (-c^t x) - \min_{x \in R \cap X} (-c^t x)}{\min_{x \in X_\rho} \varphi(x)}.$$

define

$$\mu_G := \max \Big(\frac{\| c \|}{1 - \rho}, \lambda_\rho \Big).$$

Theorem 3.1 of [2] states the equivalence between (4) and (8) for any $\mu > \mu_G$. Since μ_G depends from the radius ρ, the notation $\mu_{G,\rho}$ is used instead of μ_G.

Section 2 deals with cases in which $\mu_G \leq \mu_R$ and conditions such that $\mu_G < \mu_R$ are given.

2. Improvement of the Penalty Parameter

The starting point of the present investigation is the comparison between μ_R and λ_ρ. To this end, let us consider the following

Lemma 2.1. *Let* $\bar{x}$ *be a solution of* (6) *and* $\rho \in]0, 1[$. *If*

$$\min_{x \in X_\rho} \varphi(x) \geq \varphi(\bar{x}), \tag{10}$$

then

$$\lambda_\rho \leq \mu_R.$$

Proof. Since c has non negative components, then

$$\max_{x \in R \cap X} (-c^t x) - \min_{x \in R \cap X} (-c^t x) \leq - \min_{x \in R \cap X} (-c^t x) = c^t x_0.$$

Since $\varphi(\bar{x}) > 0$ and $\varphi(\bar{x}) = n/4 - g(\bar{x})$, then (10) implies the thesis:

$$\lambda_\rho = \frac{(\max_{x \in R \cap X} (-c^t x)) - (\min_{x \in R \cap X} (-c^t x))}{\min_{x \in X_\rho} \varphi(x)} \leq \frac{c^t x_0}{\varphi(\bar{x})} = \mu_R. \qquad \blacksquare$$

Lemma 2.2. *Let $\bar{x}$ a solution of* (6)*; if*

$$g(\bar{x}) \geq \frac{n-1}{4} \tag{11}$$

then a radius $\bar{\rho} \in]0, 1/2]$ is determined such that, for all $\rho \in [\bar{\rho}, 1[$, (10) *holds.*

Proof. (11) ensures that the intersection between the sphère

$$\Gamma = \{x \in \mathbb{R}^n : g(x) = g(\bar{x})\}$$

and any one-dimensional face of the hypercube X is non empty; whitout any loss of generality, consider the one dimensional face

$$F_1 = \{x \in \mathbb{R}^n : x = (x_1, 0, \ldots, 0),\ x_1 \in [0, 1]\},$$

and put $\bar{\rho} = d(0, F_1 \cap \Gamma)$ where 0 denotes the origin of $\mathbb{R}^n$. Now, observe that for all $x \in \{x \in \mathbb{R}^n : g > g(\bar{x})\} \cap R \cap X$, there is a $z \in Z$ such that $x \in B(z, \bar{\rho})$. The inclusion $\hat{F} \subset \{x \in \mathbb{R}^n : g \leq g(\bar{x})\}$ and the convexity of $R \cap X$ implies $z \in Z \cap R$ and then $x \notin X_{\bar{\rho}}$. Conclusion:

$$X_{\bar{\rho}} \subseteq \{x \in \mathbb{R}^n : g(x) \leq g(\bar{x})\}. \tag{12}$$

Since, for all $\rho \in [\bar{\rho}, 1[$,

$$X_\rho \subseteq X_{\bar{\rho}} \subseteq \{x \in \mathbb{R}^n : g(x) \leq g(\bar{x})\} = \{x \in \mathbb{R}^n : \varphi(x) \geq \varphi(\bar{x})\},$$

the thesis follows. ∎

In the sequel,

$$\hat{\rho} := d(\bar{x}, R \cap Z), \quad \bar{\rho} := d(0, F_1 \cap \Gamma),$$

being $F_1 = \{x = (x_1, 0, \ldots, 0) \in \mathbb{R}^n,\ x_1 \in [0, 1]\},\ \Gamma = \{x \in \mathbb{R}^n : g(x) = g(\bar{x})\}$.

Observe $\bar{\rho}$ is uniquely determined even if $\bar{x}$ is not the unique solution of (6).

The following result gives sufficient conditions ensuring it holds $\mu_{G,\bar{\rho}} \leq \mu_R$.

Theorem 2.1. *Let $\bar{x}$ be a solution of* (6) *satisfying* (11) *and let x_0 be a solution of* (5)*. If*

$$\frac{c^t x_0}{\| c \|} \geq \bar{\rho}, \tag{13}$$

then $\mu_{G,\bar{\rho}} \leq \mu_R$.

Proof. Lemma 2.2 implies

$$\lambda_{\bar{\rho}} \leq \mu_R. \tag{14}$$

Besides, since $\varphi(\bar{x}) = \bar{\rho}(1-\bar{\rho})$, thanks to (13), it holds

$$\frac{\parallel c \parallel}{1-\bar{\rho}} \leq \frac{c^t x_0}{\bar{\rho}(1-\bar{\rho})} = \mu_R. \tag{15}$$

The thesis follows from (14) and (15). ■

Theorem 2.1 assures the existence of cases in which $\mu_{G,\rho} < \mu_R$, as showed in the following corollaries.

Corollary 2.1. *Under the hypotheses of Theorem* 2.1*, let the inequality* (13) *be strict. If* $0 \notin R \cap X$*, then* $\mu_{G,\bar{\rho}} < \mu_R$.

Proof. Since $\varphi(\bar{x}) = \bar{\rho}(1-\bar{\rho})$, the inequality $c^t x_0 / \parallel c \parallel > \bar{\rho}$ is equivalent to

$$\frac{\parallel c \parallel}{1-\bar{\rho}} < \mu_R. \tag{16}$$

Moreover the hyphotesis $0 \notin R \cap X$ implies $\max_{x \in R \cap X}(-\langle c, x\rangle) < 0$, and Theorem 2.1 assures $\min_{x \in X_{\bar{\rho}}} \varphi(x) \geq \varphi(\bar{x})$. Then

$$\lambda_{\bar{\rho}} = \frac{\max_{x \in R\cap X}(-c^t x) - \min_{x \in R \cap X}(-c^t x)}{\min_{x \in X_{\bar{\rho}}} \varphi(x)} < \frac{c^t x_0}{\varphi(\bar{x})} = \mu_R. \tag{17}$$

(16) and (17) imply the thesis. ■

Corollary 2.2. *Under the hypotheses of Theorem* 2.1*, let* $\bar{x}$ *be the unique solution of* (6)*. If* $\hat{\rho} \leq \bar{\rho}$ *and*

$$\frac{\parallel c \parallel}{1-\bar{\rho}} < \lambda_{\bar{\rho}}, \tag{18}$$

then there is ρ^* *such that*

$$\forall \rho \in]\bar{\rho}, \rho^*[, \quad \mu_{G,\rho} < \mu_{G,\bar{\rho}}. \tag{19}$$

Proof. The strict inequality (18) implies there exists $\rho^* > \bar{\rho}$ such that,

$$\forall \rho \in]\bar{\rho}, \rho^*[, \quad \frac{\parallel c \parallel}{1-\bar{\rho}} < \frac{\parallel c \parallel}{1-\rho} < \frac{\parallel c \parallel}{1-\rho^*} = \lambda_{\bar{\rho}}. \tag{20}$$

To have (19) it is enough to proove that,

$$(21) \qquad \forall \rho \in]\bar{\rho}, \rho^*[, \quad \lambda_\rho < \lambda_{\bar{\rho}},$$

and this is assured if, $\forall \rho \in]\bar{\rho}, \rho^*[$, $\min_{X_\rho} \varphi(x) > \varphi(\bar{x})$. Being X_ρ a compact set this is equivalent to have $X_\rho \subseteq \{\varphi > \varphi(\bar{x})\}$ that is

$$(22) \qquad X_\rho \subseteq \{g < g(\bar{x})\}$$

Ab absurdo, let $\tilde{x} \in X_\rho \cap \{g \geq g(\bar{x})\}$. Since $\rho > \bar{\rho}$, $X_\rho \subseteq X_{\bar{\rho}} \subseteq \{g \leq g(\bar{x})\}$, then $g(\tilde{x}) = g(\bar{x})$. Moreover $\tilde{x} \neq \bar{x}$: in fact, since $\rho > \hat{\rho}$ there is $\hat{z} \in Z \cap R$ such that $\bar{x} \in B(\hat{z}, \rho)$ and then $\bar{x} \notin X_\rho$. On the contrary, since $\tilde{x} \notin X_\rho$, there is $\tilde{z} \in Z \setminus R$ such that $\tilde{x} \in B(\tilde{z}, \rho)$. Now, put

$$A = B(\tilde{z}, \rho) \cap \{g > g(\bar{x})\}$$

and observe: $\bar{x}$ is the unique solution of (6) so $A \cap \tilde{F} = \emptyset$; moreover $\hat{z} \in Z \setminus R$ and $\rho < 1$ imply $A \cap (R \cap Z) = \emptyset$. Finally, since $A \cap (conv(A^c)) = \emptyset$ it follows that $A \cap (R \cap X) = \emptyset$. This implies $\tilde{x} \in \hat{F}$, and this is absurd being $\bar{x}$ is the unique solution of (6). So (22) follows; (20) and (21) imply the thesis. ■

The following result completes the study when $\hat{\rho} > \bar{\rho}$.

Corollary 2.3. *Under the hypotheses of Theorem* 2.1*, let $\bar{x}$ be the unique solution of* (6)*. If $\bar{\rho} < \hat{\rho} < 1$ and*

$$\frac{\| c \|}{1 - \hat{\rho}} < \lambda_{\hat{\rho}}$$

then there exists ρ^* such that, $\forall \rho \in]\hat{\rho}, \rho^*[$, $\mu_{G,\rho} < \mu_{G,\hat{\rho}}$.

Proof. The proof is the same of Corollary 2.2 where $\bar{\rho}$ is replaced with $\hat{\rho}$. ■

Remark 2.1. If $\bar{x}$ is not the unique solution of (6), Corollary 2.2 and Corollary 2.3 can be generalized. If $\bar{x}^1, \ldots, \bar{x}^k$ are the solutions, it is enough to replace the amount $\hat{\rho}$ with the following one

$$\max \{d(\bar{x}^j, R \cap Z), \ j = 1, \ldots, k\}.$$

3. Concluding Remarks

The paper shows a class of problems for which the penalty parameter of [4] can be improved. In the present research the computational cost of the two methods is not analysed; indeed, none of them can be considered for solving a concrete (large-scale) problem. It is well known that no rigorous method can be used for solving efficiently such problems and that a concrete one requires an heuristic procedure. The present study aims at providing such a procedure with rigorous methods. It would be interesting to test computationally heuristic procedures utilizing the above approaches.

Moreover, observe that the request

$$g(\bar{x}) \geq \frac{n-1}{4}$$

about the solution $\bar{x}$ of (6), is verified if there is a point $\tilde{x}$ such that

$$g(\tilde{x}) \geq \frac{n-1}{4}.$$

So Theorem 2.1 don't need solve (6).

Finally, note that, differently from what has been said in $[4]$, the improvement of μ_R is possible even if the solution x_0 of (5) coincides with $\bar{x}$.

REFERENCES

1. C. Antoni and F. Giannessi, On the equivalence, via relaxation-penalization, between vector generalized systems, *Acta Matematica Vietnamica*, **22(2)** (1997), 567-588.
2. F. Giannessi and F. Niccolucci, Connections between nonlinear and integer programming problems, in: *Symposia Mathematica*, Academic Press, New York, 19, 1976, pp. 161-176.
3. F. Giannessi and F. Tardella, Connections between nonlinear programming and discrete optimization, in: *Handbook of Combinatorial Optimization*, D. Z. Du and P. M. Pardalos (eds.), Kluwer Academic Publishers, 1, 1998, pp. 149-188.
4. B. Kalantari and J. B. Rosen, Penalty for zero-one integer programming problems, *Mathematical Programming*, North Holland Publishing Company, 1982, pp. 229-232.
5. M. Ragavachari, On the connection between zero-one integer programming and concave programming under linear constraints, *Operation Research*, **17** (1969), 680-683.

Carla Antoni
Naval Academy
72 Viale Italia
Livorno, Italy
E-mail: carla.antoni5@gmail.com

Mario Pedrazzoli
Department of Mathematics
University of Pisa
Italy
E-mail: pedrazzoli.m@alice.it