Analele științifice ale Universității „Al. I. Cuza" din Iași
Tomul XXXIV, s. I a. Matematică, 1988, f. 1

ON A CLASS OF PENALTY FUNCTIONAL METHODS FOR SOLUTION OF CONVEX OPTIMIZATION PROBLEMS

BY

C. ILIOI

The method of penalty functionals is one of the simplest and widely known of the methods for solving the problem of constrained optimization, by reducing it to a set of problems of unconstrained optimization. Usually, for a minimization problem on a closed domain, the penalized functional is chosen so that it coincides with the functional to be minimized in the admissible domain and increases steeply outside it. Theoretically the convergence of this method can be proved under mild hypotheses, but in practice it is known that some difficulties arise even for convex problems. It was proved [7] that the size of the region of superlinear convergence for the minimization of the penalized functionals tends to be small and even a method which at the limit, theoretically, converges rapidly, can become ineffective by this reason. Moreover the solutions of the penalized problems do not belong to the admissible domain, so that, by such a method, we are in the unnatural position to estimate the solution of the constrained problem by approximations which do not satisfy even the constraints.

In the present paper we consider a class of penalty functionals for which admissible approximations of the solution are obtained at each step, for the convex problems. The rate of convergence is studied for the corresponding penalty methods. Also some special examples are considered.

1. Let X be a Banach reflexive space with the norm $\|.\|$, X^* the dual space with the norm denoted by $\|.\|_*$, $C\subset X$ a convex closed domain and $F: X\to]-\infty,\infty]$ a functional with the effective domain $D(F)$ such that $D(F)\cap C\neq\emptyset$. Consider the optimization problem:

(P) *find* $u\in C$ *such that* $F(u)=\inf\{F(x)\,;\,x\in C\}$.

In [4], [5], [7] by penalty functionals it is meant a set of functionals $p: X\times]0,\infty[\to[0,\infty[$ satisfying for each $\lambda>0$, the conditions:

(i) $p(x,\lambda)=0,\ \forall x\in C$;

(ii) $p(x,\lambda)>0$, $p(x,.)$ is decreasing, $\lim\limits_{\lambda\to 0} p(x,\lambda)=\infty$, $\forall x\notin C$.

The penalized functionals $F_\lambda: X\to]-\infty,\infty]$ are defined by $F_\lambda(x)=F(x)+p(x,\lambda)$ and the penalized optimization problems are the following:

(P_λ) *find* $u_\lambda\in X$ *such that* $F_\lambda(u_\lambda)=\inf\{F_\lambda(x)\,:\,x\in X\}$.

The concept of penalty functionals can be extended by substituting (i) with the weaker condition:

(i′) $$\lim_{\lambda\to 0} p(x,\lambda)=0,\quad \forall x\in C.$$

With the definition (i′), (ii) for the penalty functionals, a theorem of convergence holds, which is similar with that proven in [5].

Theorem 1. *Let $F: X\to]-\infty,\infty]$ be a continuous, strictly quasi-convex and coercive functional and let the set of penalty functionals satisfying the conditions (i′), (ii) and F_λ be lower semi-continuous and quasi-convex. Then there exist unique solutions u respectively u_λ for the problems (P) and (P$_\lambda$) and $\{u_\lambda\}$ is weakly convergent to u, $u_\lambda\xrightarrow{w} u$, when $\lambda\to 0$. Moreover $\lim\limits_{\lambda\to 0} F_\lambda(u_\lambda)=\lim\limits_{\lambda\to 0} F(u_\lambda)=F(u)$.*

If in addition the functional F is supposed to be locally uniformly quasi-convex, then the convergence of $\{u_\lambda\}$ to u is strong and the family of problems (P$_\lambda$), (P) is equiwellset.

Proof. Since F and F_λ are quasi-convex and lower semi-continuous, they are weakly lower semi-continuous and there exists unique solutions u and u_λ for the problems (P), respectively (P$_\lambda$), $\forall\lambda>0$. The condition (i′) implies:

$$\limsup_{\lambda\to 0} F_\lambda(u_\lambda)=\limsup_{\lambda\to 0} F(u_\lambda)+p(u_\lambda,\lambda))\leqslant F(u),$$

where from $\{F(u_\lambda)\}$ is upper bounded, so that $\{u_\lambda\}$ is bounded. Hence there exists the sequence $Q\{\lambda_k\}$, $\lim\limits_{k\to\infty}\lambda_k=0$, such that $u_{\lambda_k}\xrightarrow{w} u^*$. Let us suppose that u^* does not belong to C and let (eventually a subsequence) $u_{\lambda_k}\notin C$. The condition (ii) implies for each $\lambda>\lambda_k$:

$$p(u_{\lambda_k},\lambda)\leqslant p(u_{\lambda_k},\lambda_k)\leqslant F_{\lambda_k}(u_{\lambda_k})-F(u_{\lambda_k})\leqslant F_{\lambda_k}(u_{\lambda_k})-\gamma,$$

where $\gamma=\inf\{F(x);\,x\in X\}$. For $k\to\infty$, we obtain at the limit

$$p(u^*,\lambda)\leqslant \liminf_{k\to\infty} p(u_{\lambda_k},\lambda)\leqslant F(u)-\gamma,$$

which is a contradiction of the condition (ii). It follows that our supposition is not true, so that $u^*\in C$. Passing to the limit in the inequality:

(1) $$F(u_{\lambda_k})\leqslant F_{\lambda_k}(u_{\lambda_k})\leqslant F(u)+p(u,\lambda_k),$$

we get to:

$$F(u^*)\leqslant \liminf_{k\to\infty} F(u_{\lambda_k})\leqslant \limsup_{k\to\infty} F(u_{\lambda_k})\leqslant \limsup_{k\to\infty} F_{\lambda_k}(u_{\lambda_k})\leqslant F(u).$$

Since $u^*\in C$ and because of the uniqueness of the solution u of the problem (P) we have $u^*=u$. Therefore u_λ is weakly convergent to u when $\lambda\to 0$ and

$\lim\limits_{\lambda\to 0} F(u_\lambda)=F(u)$. Also from the inequality (1) we get $\lim\limits_{\lambda\to 0} F_\lambda(u_\lambda)=F(u)$ and the first part of the theorem is completely proved.

When the functional F is locally uniformly quasi-convex the proof of the second part follows the similar proof from [5].

Remark 1. The conclusions of the theorem hold if the properties (ii) of the penalty functionals are replaced by the condition:

(ii′) $$v_k\xrightarrow{w} v,\ v\notin C,\ \lambda_k\to 0 \text{ imply } \lim_{k\to\infty} p(v_k,\lambda_k)=\infty.$$

(which is used in [3] for finite dimensional problems).

Remark 2. The condition (ii) or (ii′) could be removed, whether it is provided that the solutions u_λ of the penalized problems (P$_\lambda$) belong to the domain C for any $\lambda>0$. Moreover in this case we would obtain admissible approximations for the solution u of the problem (P).

2. Let us consider now the domain C of the form

(2) $$C=\{x\in X;\ G(x)\leqslant 0\},$$

with $G: X\to]-\infty,\infty]$, satisfying a Slater condition

(3) $$\exists v\in X,\ G(v)<0.$$

Let us consider a convex differentiable function $\varphi: [0,\infty[\to[0,\infty[$ satisfying the conditions

(4) $$\varphi(0)=\varphi'(0)=0;\ \exists\alpha=\text{const.},\ \varphi'(\lambda)/\lambda\geqslant\alpha>0,\ \forall\lambda\geqslant 0.$$

In the sequel we study the penalty functionals of the form

(5) $$p(x,\lambda)=c\varphi((G(x)+\lambda)_+)/\lambda$$

where c is a constant.

Theorem 2. *Let $F: X\to]-\infty,\infty]$ be convex differentiable and satisfying the hypotheses of Theorem 1, and $G: X\to]-\infty,\infty]$ convex differentiable such that the following hypothesis holds:*

(6) $$\sup\{\|F'(x)\|_*/\|G'(x)\|_*;\ x\in X,\ G(x)=0\}=\beta<\infty.$$

Then, by taking the penalty functionals of the form (5) where the function φ satisfies the conditions (4) and under the hypothesis

(7) $$c>\beta/\alpha,$$

the conclusions of Theorem 1 are valid and all the solutions u_λ of the problems (P$_\lambda$) belong to the admissible domain C.

Proof. For each $x\in C$ the functionals p satisfy the condition (i′), so as in order to prove the theorem we need only to verify that the solutions u_λ of the penalized problems belong to the domain C. Let us suppose that there exists $\lambda>0$ such that $u_\lambda\notin C$ and consider $z\in C$ so that $G(z)=0$. Then

(8) $$\langle F_\lambda'(z),\ z-u_\lambda\rangle=\langle F'(z),\ z-u_\lambda\rangle+c\varphi'(\lambda)\langle G'(z),\ z-u_\lambda\rangle/\lambda.$$

Let z be the best approximation of the solution u_λ on the domain C, which exists in the reflexive space X. Then z is a minimizing point for the functional $g(x)=\|x-u_\lambda\|^2/2$ on the set C, that is $0\in\partial g(z)+\partial I_C(z)$, where I_C is the indicator function of C. Let $z^*\in\partial g(z)\cap-\partial I_C(z)$. As $z^*\in\partial g(z)$ the equality $\langle z^*, z-u_\lambda\rangle=\|z^*\|_*\|z^*\|_*\|z-u_\lambda\|$ holds [1]. Since $-z^*\in\partial I_C(z)$, $-z^*$ belongs to the normal cone of C at z, which is the cone spanned by $\partial G(z)=\{G'(z)\}$, [6]. Therefore there exists $\mu>0$ such that $-z^*=\mu G'(z)$ and thus $\langle G'(z), z-u_\lambda\rangle=-\|G'(z)\|_*\|z-u_\lambda\|$. So, from the equality (8), we obtain

$$\langle F_\lambda'(z), z-u_\lambda\rangle=\langle F'(z), z-u_\lambda\rangle-c\varphi'(\lambda)\|G'(z)\|_*\|z-u_\lambda\|/\lambda\leqslant$$

$$\leqslant(\|F'(z)\|_*-c\varphi'(\lambda)\|G'(z)\|_*/\lambda)\|z-u_\lambda\|.$$

By using the hypothesis (7) and the properties (4) of the function φ, we get the inequality $\langle F_\lambda'(z), z-u_\lambda\rangle<0$. On the other side, since $F'_\lambda(u_\lambda)=0$ and because of the convexity of the functional F_λ, it follows

$$\langle F_\lambda'(z), z-u_\lambda\rangle=\langle F_\lambda'(z)-F_\lambda'(u_\lambda), z-u_\lambda\rangle\geqslant 0. \tag{9}$$

This contradiction proves that, for each $\lambda>0$, the solutions u_λ belong to the domain C. The theorem is completely proved.

Consequence. *Under the hypotheses of the theorem 2 the rate of convergence of the optimal values of the penalized problems* (P_λ) *to the optimal value of the problem* (P) *is given by the inequalities*:

$$F(u)\leqslant F_\lambda(u_\lambda)\leqslant F(u)+c\varphi(\lambda)/\lambda. \tag{10}$$

Proof. Since we proved that the solutions u_λ of the problems (P_λ) belong to the domain C, we can write:

$$F_\lambda(u_\lambda)\geqslant F(u)+p(u_\lambda, \lambda)\geqslant F(u).$$

Also, by using the inequality $F_\lambda(u_\lambda)\leqslant F_\lambda(u)$, we obtain the second relation from (10).

This estimation implies a rate of convergence depending on the convergence to zero of $\varphi(\lambda)/\lambda$ when $\lambda\to 0$. For example when $\varphi(\lambda)=\lambda^2$ we get a rate of convergence of the order $O(\lambda)$. By using a superlinear method for the approximation of u at each step, we obtain an efficient penalty method.

3. We consider now an example of optimization problem with operatorial constraints, for which the same type of penalty functionals is applied. In this case we shall see that the condition as the solutions of the penalized problems to belong to the admissible domain is much simpler than that of the Theorem 2.

Let us consider a problem of the form (P) — an obstacle problem — with [2]:

$$F(v)=\tfrac{1}{2}\int_\Omega|\operatorname{grad} v|^2\,dx-\int_\Omega f(x)v(x)dx \tag{11}$$

and

$$C=\{v\,;\ v\in H_0^1(\Omega),\ v(x)\geqslant\delta(x)\ \text{a.e.}\}. \tag{12}$$

We denote here by Ω a bounded open domain of R^n with the boundary Γ, a $(n-1)$-dimensional variety of class C^∞, and by $\delta\in H^1(\Omega)$ the obstacle function. We suppose that $F'(\delta)\in L^\infty(\bar\Omega)$ and $c>0$ is a constant satisfying the inequality:

$$c\geqslant\|F'(\delta)\|_\infty. \tag{13}$$

For solving this problem by penalty functional method we take:

$$p(v, \lambda)=c\|(\delta-v+\lambda)_+\|^2/2\lambda, \tag{14}$$

where the norm is that of $L^2(\Omega)$. We shall prove that the solutions of the penalized problems belong to the set C, when the inequality (13) is satisfied. Let us suppose that there exists a $\lambda>0$ such that the solution u_λ of (P_λ) does not belong to C. Then a function $v_\lambda\neq u_\lambda$ can be defined by

$$v_\lambda(x)=\begin{cases}u_\lambda(x) & \text{when } u_\lambda(x)\geqslant\delta(x),\\ \delta(x) & \text{otherwise}\end{cases} \tag{15}$$

Since the functional F is strictly convex it follows that

$$(F'(v_\lambda)-F'(u_\lambda),\ v_\lambda-u_\lambda)>0, \tag{16}$$

where by $(\ldots)$ we denote the inner product of $L^2(\Omega)$. Then, since u_λ is the solution of the unconstrained problem (P_λ), we have

$$0=F_\lambda'(u_\lambda)=F'(u_\lambda)+p'(u_\lambda, \lambda)$$

and because of the convexity of $p'(\cdot, \lambda)$ we can write successively:

$$0<(F'(v_\lambda)+p'(u_\lambda, \lambda),\ v_\lambda-u_\lambda)\leqslant(F'(v_\lambda)+p'(v_\lambda, \lambda),\ v_\lambda-u_\lambda)=$$

$$=\int_\Omega\{\operatorname{grad} v_\lambda(x)\cdot\operatorname{grad}(v_\lambda(x)-u_\lambda(x))-f(x)(v_\lambda(x)-u_\lambda(x))\,dx-$$

$$-c\int_\Omega\{\delta(x)-v_\lambda(x)+\lambda)_+(v_\lambda(x)-u_\lambda(x))/\lambda\}dx.$$

Let us denote

$$\Omega_1=\{x\in\Omega\,;\ v_\lambda(x)=\delta(x)\}=\mathrm{cl}\{x\in\Omega\,;\ \delta(x)>u_\lambda(x)\}.$$

Then we have the inequality

$$0<\int_{\Omega_1}\{\operatorname{grad} v_\lambda(x)\cdot\operatorname{grad}(\delta(x)-u_\lambda(x))-f(x)(\delta(x)-u_\lambda(x))\}dx-c\int_{\Omega_1}(\delta(x)-$$

$$-u_\lambda(x))dx.$$

where from

$$0<\int_{\Omega_1}(F'(\delta(x))-c)(\delta(x)-u_\lambda(x))dx\leqslant(\|F'(\delta)\|_\infty-c)\int_{\Omega_1}(\delta(x)-u_\lambda(x))dx. \tag{17}$$

But this is in contradiction with the hypothesis (13). So all the solutions u_λ belong to the set C and the Theorem 1 can be applied.

REFERENCES

1. Barbu V., Precupanu T. — *Convexity and Optimization in Banach Spaces*, Editura Academiei R.S.R., București, D. Reidel Publ. Co. Dordrecht, Boston, Lancaster, 1986.
2. Glowinski R., Lions J. L., Trémolières R. — *Analyse Numérique des Inéquations Variationelles*, Dunod, Paris, 1976.
3. Grossman K., Kaplan A. A. — *Nonlinear Programming on the Basis of Unconstrained Minimization* (in Russian), Nauka, Novosibirsk, 1981
4. Ililoi C. — *Optimization Problems and Algorithms for Approximate Solutions* (in Roumanian), Editura Academiei R. S. R., 1980.
5. Ililoi C. — *On the locally uniformly quasi convex functionals and the penalti functional methods*. Boll. U. M. I., (6), 4B, 1986, 1 360–1 370.
6. Ioffe A. D., Tihomirov V. M. — *Theory of the Extremal Problems* (in Russian), Nauka, Moskow, 1974.
7. Pshenichnyi B. N., Danilin Yu. M. — *Numerical Methods in Extremal Problems*, Mir Publisher, Moskow, 1978.

Received 26.XI. 1986

Faculty of Mathematics
University of Iași
6600 — Iași, România

Analele științifice ale Universității „Al. I. Cuza" din Iași
Tomul XXXIV, s. I a. Matematică, 1988, f. 1

MAXIMUM PRINCIPLE FOR CONTROL PROBLEMS GOVERNED BY BILINEAR FINITE DIMENSIONAL INEQUALITIES

BY

CĂTĂLINA DAVIDEANU

This paper is concerned with necessary optimality conditions (maximum principle) for control problems governed by bilinear finite dimensional inequalities.

A simple example is also presented, in which the optimality conditions established are applied and the optimal pair is found.

1. Necessary optimality conditions. We consider the following control problem :

Problem (P_K) : *Minimize*

$$\int_0^T [g(t, y(t)) + h(u(t))]\mathrm{d}t + \varphi_0(y(T)),\quad 0 < T < \infty, \tag{1.1}$$

over the set of all functions $y \in W^{1,2}([0, T] ; R^n)$ and $u \in L^2(0, T)$ *subject to* :

$$y'(t) + (F + \partial I_K)(y(t)) + u(t)F_1(y(t)) \ni f(t) \quad a.e.\ t \in]0, T[, \tag{1.2}$$

$$y(0) = y_0 \in K, \tag{1.3}$$

where $K = \{z \in R^n ;\ 0 \leqslant z_i \leqslant a_i,\ i = 1, 2, \ldots, n\}$.

Using the definition of the operator ∂I_K, the relations (1.2) and (1.3) can be equivalently written as :

$$\begin{aligned} &y_i'(t) + (Fy(t))_i + u(t)(F_1y(t))_i = f_i(t) \ \text{a.e.}\ \{t \in [0, T] ;\ 0 < y_i(t) < a_i\},\\ &y_i'(t) + (Fy(t))_i + u(t)(F_1y(t))_i \geqslant f_i(t) \ \text{a.e.}\ \{t \in [0, T] ;\ y_i(t) = 0\},\\ &y_i'(t) + (Fy(t))_i + u(t)(F_1y(t))_i \leqslant f_i(t) \ \text{a.e.}\ \{t \in [0, T] ;\ y_i(t) = a_i\},\\ &0 \leqslant y_i(t) \leqslant a_i,\ \text{a.e.}\ t \in]0, T[,\ y_i(0) = y_{0,i} \in [0, a_i], \end{aligned} \tag{1.4}$$

where $i = 1, 2, \ldots, n$.

The elements which have appeared satisfy the following properties :

(i) $F : R^n \to R^n$ is locally Lipschitz and there exist $D_1, D_2 \in R_+$ such that

$$(Fy, y) \geqslant -D_1 \| y \|^2 - D_2 \quad \text{for all } y \in R^n.$$

(ii) $F_1 : R^n \to R^n$ is a M-Lipschitzian function.