INSTITUTUL DE MATEMATICA
AL ACADEMIEI ROMANE

PREPRINT SERIES OF THE INSTITUTE OF MATHEMATICS
OF THE ROMANIAN ACADEMY

ISSN 0250 3638

ON THE CONVEX PROGRAMMING PROBLEM

BY
DAN TIBA

Preprint No. 2/1995

BUCURESTI

ON THE CONVEX PROGRAMMING PROBLEM

BY

DAN TIBA*)

January 1995

*) Institute of Mathematics of the Romanian Academy, P.O.Box 1-764, RO-70700 Bucharest, Romania

On the convex programming problem

by
Dan Tiba
Institute of Mathematics
Romanian Academy
P.O.Box 1–764
RO–70700 Bucharest, ROMANIA
E–mail address: dtiba@imar.ro

Abstract

The nondegenerate Lagrange multipliers rule is proved under an explicit constraint qualification strictly weaker than the Slater condition. A first approach is based on direct exact penalization arguments, while a second one uses the theory of maximal monotone operators and allows a large class of applications.

Mathematics Subject Classification: 49B27

January 1995

1. Introduction

Let H be a Banach space, $g : H \to]-\infty, +\infty]$ be a convex, lower semicontinuous proper function and $h_i : H \to]-\infty, +\infty]$, $i = \overline{1,n}$, be convex mappings. Precise hypotheses will be stated in the sequel, for each result. We consider the standard convex programming problem with inequality constraints:

$$\text{(1.1)} \qquad \text{Minimize } \{g(x)\}$$

subject to

$$\text{(1.2)} \qquad h_i(x) \leq 0, \; i = \overline{1,n}.$$

We define the feasible (convex) set

$$\text{(1.3)} \qquad C = \{x \in H;\; h_i(x) \leq 0, \; i = \overline{1,n}\}$$

assumed to be nonvoid (admissibility) and we denote by $\overline{x} \in C$ a solution of (1.1), (1.2), supposed to exist. We also introduce the convex mapping $h : H \to]-\infty, +\infty]$, $h(x) = \max\{h_i(x);\; i = \overline{1,n}\}$ and the problem (1.1), (1.2) may be equivalently reformulated as (1.1) and

$$\text{(1.4)} \qquad h(x) \leq 0.$$

This work discusses the classical question of the Lagrange multipliers associated to $\overline{x}$ and of the related constraint qualification conditions if a "nondegenerate" (involving g essentially) characterization is obtained.

In the next section, we show by a direct argument a geometric property of convex functions (valid on $H \backslash C$) which is a consequence of the Slater [11] condition. This allows to obtain a modified nondegenerate Lagrange multipliers rule involving h_+, the positive part of h.

A new explicit constraint qualification, strictly weaker than the Slater condition, is formulated and the standard Karush-Kuhn-Tucker [7], [8] conditions are reobtained (under this assumption) via subdifferential calculus.

In the last section the continuity hypotheses on g are relaxed by means of an alternative approach based on the Minty [9] theorem on maximal monotone operators in Banach spaces, Barbu [2].

Finally, we underline that the literature on the relaxation of constraint qualifications in the mathematical programming or of interiority conditions in optimal control is quite rich: Zowe and Kurcyusz [17], Clarke [5, Ch. VI], Troltzch [16], Barbu and Pavel [3], Tiba [12], Neittaanmaki and Tiba [13, Ch. VI], Tiba and Bergounioux [14], Vinter and Ferreira [18]. Generally, qualification conditions on the feasible set are required, while our hypothesis is related to the behaviour of h on the complementary of C.

2. Slater condition and exact penalization

Theorem 2.1. *Let h be convex proper lower semicontinuous and let the Slater assumption*

$$\exists \hat{x} \in C : h(\hat{x}) < 0 \tag{2.1}$$

be fulfilled. Then, $\forall\, r > 0$, $\forall\, x \in B(\hat{x}, r) \setminus C$, *we have*

$$h(x) \geq -\frac{h(\hat{x})}{r} \operatorname{dist}(x, C). \tag{2.2}$$

Proof.

Take $a = -h(\hat{x}) > 0$ and denote by S the line passing through $\hat{x}, x$ in H. If $h(x) = +\infty$, then (2.2) is trivial. Therefore, we may assume that h is finite on the closed segment $[\hat{x}x] \subset S$ since it is convex, proper. Then $h|_S$ is continuous on the open segment $]\hat{x}x[\subset S$. As $x \in H \setminus C$ and C is closed, there is $\varepsilon > 0$ such that $h > 0$ on $B(x, \varepsilon) \cap]\hat{x}x[$ due to (1.3). Since h is convex, it follows that $\lim_{y \to x-} h(y)|_S > 0$.

By, Proposition 3.1.2., Hiriart–Urruty and Lemarechal [6, Ch. I], we have that

$$\lim_{y \to \hat{x}+} h(y)|_S \leq h(\hat{x}) < 0.$$

We suppose that a unit vector u, $|u|_H = 1$ is chosen parallel to S and a parametrization of S with respect to u and some origin is given. The above discussion shows the existence of $\tilde{x} \in]\hat{x}x[$ such that $h(\tilde{x}) = 0$, $\tilde{x} \notin C$. Let $\hat{\lambda}$, $\tilde{\lambda}$, λ be the "coordinates" of $\hat{x}$, $\tilde{x}$, x on S, respectively and assume that $\hat{\lambda} < \tilde{\lambda} < \lambda$.

For $y \in S$, $y \cong \mu u$, $\mu \in R$, we define

$$f(y) = \frac{a}{r}(\mu - \tilde{\lambda}) \tag{2.3}$$

f being an affine mapping on S. We notice by (2.3) that

$$f(\tilde{x}) = 0 = h(\tilde{x}), \tag{2.4}$$

$$f(\hat{x}) = -\frac{h(\hat{x})}{r}(\hat{\lambda} - \tilde{\lambda}) = h(\hat{x})\frac{\tilde{\lambda} - \hat{\lambda}}{r} \geq h(\hat{x}) \tag{2.5}$$

since $h(\hat{x}) < 0$ and $0 \leq \dfrac{\tilde{\lambda} - \hat{\lambda}}{r} = \dfrac{|\, \tilde{x} - \hat{x}\,|_H}{r} \leq 1$. Due to the convexity of $h|_S$ and the affine character of f, we get by (2.4), (2.5) that

$$h(x) \geq f(x) = \frac{a}{r}(\lambda - \tilde{\lambda}) = \frac{a}{r}\,|\, x - \tilde{x}\,|_H \geq$$

$$\geq \frac{a}{r}\,|\, x - \operatorname{proj}_C x\,|_H = -\frac{h(\hat{x})}{r} \operatorname{dist}(x, C)$$

since $\tilde{x} \in C$. This ends the proof.

Remark. Suggested by Theorem 2.1, we formulate the following explicit constraint qualification:

$$\forall\, M \subset H \text{ bounded}: \; M \setminus C \neq \emptyset, \tag{2.6}$$

$$\exists c_M > 0 : h(x) \geq c_M \mid x - \text{proj}_C x \mid_H, \; \forall\, x \in M \setminus C,$$

which is strictly weaker than (2.1) by the above result and by the example of h satisfying (2.1) and its positive part h_+ for which (2.1) fails but not (2.6).

If C is bounded, then c_M may be choosen independent of M. Moreover, in (2.6) a neighbourhood of C may be taken into account instead of H, by the convexity of h.

Remark. By (2.6) and (1.3), we have

$$h(x) - h(\text{proj}_C\, x) \geq c_M \mid x - \text{proj}_C\, x \mid_H \tag{2.6'}$$

for $x \in M \setminus C$. Relation (2.6)' expresses that the subgradients of h in x, $\partial h(x)$, are "far" from zero. This will play an essential role in the next section.

Theorem 2.2. *Let $h : H \to]-\infty, +\infty]$ be convex lower semicontinuous proper and $g : H \to R$ be convex continuous. Then, if $\overline{x}$ is a solution of (1.1), (1.4) and (2.6) is satisfied, there is $\lambda \geq 0$ such that $\overline{x}$ is a minimum point of $g + \lambda h_+$ over H.*

Proof.

Let $B(\overline{x}, \varepsilon)$ be a "small" ball around $\overline{x}$. We show the minimum property of $g + \lambda h_+$ (for some $\lambda \geq 0$) on $B(\overline{x}, \varepsilon)$ and it will follow on H, by convexity. For any $y \in B(\overline{x}, \varepsilon)$, $\text{proj}_C\, y \in B(\overline{x}, \varepsilon)$ since $\overline{x} = \text{proj}_C\, \overline{x}$ and $\text{proj}_C\,(\cdot)$ is nonexpansive. We have

$$g(y) - g(\text{proj}_C\,(y)) \geq (w, y - \text{proj}_C\, y)_{H^* \times H} \geq -L \mid y - \text{proj}_C y \mid_H$$

for any $w \in \partial g\,(\text{proj}_C\, y)$ and for $L \geq\mid w \mid_{H^*}$ given by the boundedness of $\partial g(\cdot)$ on $B(\overline{x}, \varepsilon)$ due to the continuity of g. Then

$$g(y) + L \mid y - \text{proj}_C\, y \mid_H \geq g(\text{proj}_C\, y) \geq g(\overline{x}). \tag{2.7}$$

By (2.6), we have

$$h_+(x) \geq c_\varepsilon \mid x - \text{proj}_C\, x \mid_h, \; \forall\, x \in B(\overline{x}, \varepsilon) \setminus C. \tag{2.8}$$

Then, (2.7), (2.8) yield, for $y \in B(\overline{x}, \varepsilon) \setminus C$:

$$g(y) + \frac{L}{c_\varepsilon} h_+(y) \geq g(y) + L \mid y - \text{proj}_C\, y \mid \geq \tag{2.9}$$

$$\geq g(\overline{x}) = g(\overline{x}) + \frac{L}{c_\varepsilon} h_+(\overline{x}).$$

Relation (2.9) remains valid for $y \in B(\overline{x}, \varepsilon) \cap C$ since $h_+|_C = 0$ and this finishes the proof with $\lambda = \dfrac{L}{c_\varepsilon} \geq 0$.

Remark. The above proof is based on direct exact penalization arguments as in Lemarechal and Hiriart–Urruty [6, Ch. VII 1.2]. We remark that only g locally Lipschitzian is necessary in this setting. On the contrary, based on subdifferential calculus rules (which are known to be equivalent with the separation theorem, Tichomirov [15], p. 52), we shall get the standard Karush–Kuhn–Tucker optimality conditions (Theorem 2.3 below).

Example. For $H = R$, if (2.6) fails, then the conclusion of Theorem 2.2 may be not valid. We take $\tilde{g}(x) = x$, $x \in R$ and

$$\tilde{h}(x) = \begin{cases} x^2, & x < 0 \\ 0, & x \geq 0 \end{cases}$$

and we consider the minimization of $\tilde{g}$ subject to $\tilde{h}(x) \leq 0$ (i.e. $x \geq 0$). The solution is $\overline{x} = 0$ and (2.6) is not fulfilled since $\tilde{h}' \equiv 0$ for $x \in C$. For any $\lambda \geq 0$, we define the nondegenerate Lagrange function $g_\lambda(x) = \tilde{g}(x) + \lambda \tilde{h}_+(x) = \tilde{g}(x) + \lambda \tilde{h}(x)$. Obviously, $x_\lambda = -\dfrac{1}{2\lambda}$ provides a global minimum for g_λ and $x_\lambda \neq \overline{x}$, $\forall\ \lambda > 0$. For $\lambda = 0$, the infimum on R is not attained.

Theorem 2.3. *Assume* (2.6) *and that g and h_i, $i = \overline{1,n}$, are convex continuous on H. If $\overline{x}$ is a solution of* (1.1), (1.2) *there are $\lambda_i \geq 0$, $i = \overline{1,n}$, such that*

$$0 \in \partial g(\overline{x}) + \sum_{i=1}^{n} \lambda_i \partial h_i(\overline{x}), \tag{2.10}$$

$$\lambda_i h_i(\overline{x}) = 0,\ i = \overline{1,n} \tag{2.11}$$

Proof.

This is a consequence of Theorem 2.2 and of the Dubovitskij–Milyutin theorem which is valid under the continuity assumptions, Tichomirov [15], p.52:

$$\partial h(x) = \{ \sum_{j \in T(x)} \lambda_j w_j;\ \lambda_j \in R_+,\ \sum_{j \in T(x)} \lambda_j = 1,\ w_j \in \partial h_j(x)\},$$

$$T(x) = \{j;\ h_j(x) = h(x)\}.$$

We also use that $h_+ = \sup(0, h)$ and the additivity rule $\partial(g + \lambda h_+) = \partial g + \lambda \partial h_+$ to finish the proof.

Example. As a byproduct of Theorem 2.1, we put into evidence a class of functionals on H which have the exact penalization property for convex closed sets $C \subset H$ such that $0 \in C$. We denote by $p_c : H \to]-\infty, +\infty]$ the Minkowski (gauge) functional associated to C and we define, for $\varepsilon > 0$:

$$h^\varepsilon(x) = \varepsilon(p_c(x) - 1)_+. \tag{2.10}$$

The following properties are obvious:

$$C = \{x \in H;\ h^\varepsilon(x) = 0\},\ \ \forall\, \varepsilon > 0,$$
$$h^\varepsilon(x) \to +\infty, \text{ for } \varepsilon \to \infty,\ \forall\, x \in H \setminus C,$$
$$h^\varepsilon(x) \to 0 \text{ for } \varepsilon \to 0,\ \forall\, x \in \text{ aff }(C).$$

Moreover, if $\tilde{h}^\varepsilon(x) = \varepsilon(p_c(x) - 1)$, then it satisfies the Slater condition $\tilde{h}^\varepsilon(0) = -\varepsilon$, $\forall\, \varepsilon > 0$ and Theorem 2.1 shows that h^ε satisfies (2.6). If g is a continuous convex mapping with a minimum in $\overline{x} \in C$, then (by the proof of Theorem 2.2) there is $\varepsilon > 0$ such that $\overline{x}$ is a minimum for $g + h^\varepsilon$ on H.

It is well known that the distance function, dist (x, C), has the exact penalization property, Hiriart–Urruty and Lemarechal [6, Ch. VII, 1.2], while (2.10) is an example of a different nature.

3. Maximal monotonicity

In this section we assume that H is a reflexive Banach space. Then, an equivalent norm may be defined such that H and the dual H^* are strict convex, Asplund [1].

By using a different technique, we prove the nondegenerate Lagrange multipliers rule under condition (2.6) and with general convex lower semicontinuous proper mapping $g : H \to]-\infty, +\infty]$. This is important since it allows to add to the problem (1.1), (1.2) some abstract type constraints $x \in A \subset H$ convex, closed nonvoid subset, by the standard trick of redefining g as $+\infty$ outside A.

Theorem 3.1. *Let $h : H \to R$ be convex continuous satisfying* (2.6) *and* $0 \in \partial h(0)$. *Then, the operator $N \subset H \times H^*$ given by*

$$N(x) = \begin{cases} 0, & \text{if } h(x) < 0 \\ \lambda w, \lambda \geq 0, w \in \partial h(x), & \text{if } h(x) = 0 \\ \emptyset & \text{if } h(x) > 0 \end{cases} \tag{3.1}$$

is maximal monotone and $N(x) = \partial I_C(x)$, the normal cone to C at x.

In (3.1), it is possible that no $x \in H$ satisfies $h(x) < 0$ and, then, the first line disappears.

Proof

By the definition of the subdifferential, we have $N(x) \subset N_C(x)$, $\forall\, x \in H$. This is obvious if $x \notin C$ or $h(x) < 0$. If $x \in C$ and $h(x) = 0$, we have

$$h(x) - h(y) \leq (w, x - y)_{H \times H^*},\ \forall\, x \in \partial h(x),\ \forall\, y \in H$$

that is

$$0 \leq \lambda(w, x - y)_{H \times H^*},\ \forall\, \lambda \geq 0,\ \forall\, y \in C \tag{3.2}$$

since $h(x) - h(y) = -h(y) \geq 0$. Relation (3.2) proves the inclusion. It yields that $N \subset H \times H^*$ is monotone and we have to show its maximality to get the desired equality. Since $\partial h \subset H \times H^*$ is maximal monotone, the Minty theorem, Barbu [2, Ch. II] gives the existence of $x_\lambda \in H$ such that

$$F(x_\lambda - y) + \lambda \partial h(x_\lambda) \ni 0 \tag{3.3}$$

where $\lambda > 0$, $F : H \to H^*$ is the duality mapping and $y \in H$ is arbitrary.

We show that $x_\lambda \in C$ for λ "sufficiently big". If $y \in C$, we may take $x_0 = x_\lambda = y$ and $\lambda = 0$. Then (3.3) becomes (by using (3.1)):

$$F(x_0 - y) + N(x_0) \ni 0. \tag{3.4}$$

Therefore, we may assume $h(y) > 0$. By the definition and properties of the Moreau-Yosida regularization, Barbu and Precupanu [4, Ch. 2.3], we have

$$h(x_\lambda) = h_\lambda(y) - \frac{\lambda}{2}\,|\partial h_x(y)|^2_{H^*} = \tag{3.5}$$

$$= h_\lambda(y) - \frac{\lambda}{2}\,|w_\lambda|^2_{H^*}$$

where w_λ is some element in $\partial h(x_\lambda)$.

If $x_\lambda \notin C$ for some $\lambda > 0$, then $h(x_\lambda) > 0$ and (3.5) gives

$$|w_\lambda|^2_{H^*} < \frac{2h_\lambda(y)}{\lambda} \le \frac{2h(y)}{\lambda}. \tag{3.6}$$

Taking into account hypothesis (2.6), we get:

$$c_*\,|x_\lambda - \operatorname{proj}_C x_\lambda|_H \le h(x_\lambda) - h(\operatorname{proj}_C x_\lambda) \le \tag{3.7}$$

$$\le (\partial h(x_\lambda), x_\lambda - \operatorname{proj}_C x_\lambda)_{H\times H^*} \le |\partial h(x_\lambda)|_{H^*} \cdot |x_\lambda - \operatorname{proj}_C x_\lambda|_H\,,$$

where $c_* > 0$ is a constant depending on y^* since $|x_\lambda|_H \le |y^*|_{H^*}$ by $0 \in \partial h(0)$. Combining (3.6), (3.7) we get:

$$c_*^2 < \frac{2h(y)}{\lambda}$$

which is a contradiction for λ big enough.

By this discussion and since $N(x)$ is a cone, we conclude that for any $y \in H$, the equation

$$F(x_* - y) + N(x_*) \ni 0 \tag{3.8}$$

has a solution $x_* \in C$.

Assume now that $N \subset H \times H^*$ is not maximal monotone. This means that there is $x' \in H$, $y' \in H^*$ such that $y' \notin N(x')$ and

$$(x - x',\ y - y')_{H\times H^*} \ge 0 \tag{3.9}$$

for any $x \in C$, any $y \in N(x)$.

By (3.8) there is $x^0 \in C$, solution for the equation

$$F(x^0 - x' - F^{-1}(y')) + y^0 = 0 \tag{3.10}$$

with some $y^0 \in N(x^0)$. Choosing $x = x^0$, $y = y^0$ in (3.9), we get

$$(x^0 - x', \ -F(x^0 - x' - F^{-1}(y')) - y')_{H \times H^*} \geq 0$$

that is

$$0 \geq (x^0 - x' - F^{-1}(y') - (-F^{-1}(y')), \ F(x^0 - x' - F^{-1}(y')) - F(-F^{-1}(y'))_{H \times H^*}.$$

Since F is strictly monotone in $H \times H'$, we get $x^0 = x'$ and $y^0 = y'$ by (3.10) and the proof is finished.

Remark. The inclusion $N \subset N_C$ is well known and the equality $N = N_C$ is an abstract regularity condition, necessary and sufficient for the nondegenerate Lagrange multipliers rule to hold (Hiriart-Urruty and Lemarechal [6, Ch.VII.2]). Another "basic constraint qualification" may be formulated via tangent cones as well, Rockafellar [10].

Corollary 3.2. *Let $g : H \to]-\infty, +\infty]$ be convex lower semicontinuous proper, continuous in some point of C and $\overline{x}$ be a solution of* (1.1), (1.2). *If h satisfies the assumptions of Theorem 3.1, then there are $\lambda_i \geq 0$ such that*

$$0 \in \partial g(\overline{x}) + \sum_{i=1}^{n} \lambda_i \partial h_i(\overline{x}), \ \lambda_i h_i(\overline{x}) = 0, \ i = \overline{1,n}.$$

Proof

Since g is continuous in a point of C, we have the additivity rule $0 \in \partial g(\overline{x}) + N_C(\overline{x})$. The proof is finished by Theorem 3.1 and the Dubovitskij–Milyutin theorem.

Remark. The hypothesis $0 \in \partial h(0)$ is not restrictive since if the Slater condition does not hold, then $\forall\, x \in C$ is a minimum point for h by $C = \{x \in H; \ h(x) = 0\}$ and $h \geq 0$ on H. Therefore a simple shifting on h gives $0 \in \partial h(0)$. If the Slater condition is fulfilled the results are wellknown.

References

[1] E. Asplund - "Averaged norms", Israel J. Math. 5(1967).

[2] V. Barbu - "Nonlinear semigroups and differential equations in Banach spaces", Noordhoff, Leyden (1976).

[3] V. Barbu, N. Pavel - "Optimal control problems with two point boundary condition", J.O.T.A. vol. 77 (1993).

[4] V. Barbu, Th. Precupanu - "Convexity and optimization in Banach spaces", Noordhoff, Leyden (1978).

[5] F.H.Clarke - "Optimization and nonsmooth analysis", Wiley, New York (1983).

[6] J.-B. Hiriart-Urruty, Cl. Lemarechal - "Convex analysis and minimization algorithms", Springer Verlag, Berlin (1993).

[7] W. Karush - "Minima of functions of several variables with inequalities as side conditions", Master's Thesis, Dept. of Mathematics, Chicago University (1939).

[8] H.W.Kuhn, A.W.Tucker - "Nonlinear programming", in J. Neyman (ed.), Proceedings of second Berkeley symposium on mathematical statistics and probability, Berkeley, University of California (1950).

[9] G.J.Minty - "Monotone (nonlinear) operators in Hilbert spaces", Duke Math. J. 29(1962).

[10] R.T.Rockafellar - "Lagrange multipliers in optimization", SIAM-AMS Proceedings 9 (1976).

[11] M. Slater - "Lagrange multipliers revisited", Cowles comission discussion paper no. 403, (1950).

[12] D. Tiba - "Nonqualified convex control problems", Report 424, Technische Univ. München, Inst. for Appl. Math. and Statistics (1993).

[13] D. Tiba, P. Neittaanmaki - "Optimal control of nonlinear parabolic systems", Marcel Dekker, New York (1994).

[14] D. Tiba, M. Bergounioux - "General optimality conditions for constrained convex control problems", SIAM J. Control Optimiz. (1995), (accepted).

[15] V.M. Tichomirov - "Convex analysis", Encyclopaedia of mathematical sciences vol. 14, Springer Verlag (1990).

[16] F. Tröltzsch - "Optimality conditions for parabolic control problems and applications", Teubner Texte 62, Leipzig (1984).

[17] J. Zowe, S. Kurcyusz - "Regularity and stability for the mathematical programming problem in Banach spaces", Appl. Math. Optimiz. 5 (1979).

[18] R.B. Vinter, M. Ferreira - "When is the maximum principle for state constrained problems nondegenerate", J. Math. An. Appl. (1995), (to appear).