Rigidity of Abnormal Extrema in the Problem of Non-linear Programming with Mixed Constraints

A.V.Sarychev*, T.V.Tchemisova †

Abstract

We study abnormal extremum in the problem of non-linear programming with mixed constraints. Abnormal extremum occurs when in necessary optimality conditions the Lagrange multiplier, which corresponds to the objective function, vanishes. We demonstrate that in this case abnormal second-order sufficient optimality conditions guarantee rigidity of the corresponding extremal point, which means isolatedness of this point in the set determined by the constraints.

Key words: nonlinear programming, optimality conditions, abnormal extremum, rigidity

1 Introduction

Although local optimality for smooth problems of nonlinear programming (NLP) has been extensively studied and by now is textbook material (see e.g. [7]), there exists a particular case - called abnormal extremum - which has been almost ignored till recent time. It occurs if in the classical optimality conditions (for example, in Fritz-John formulation [5]) the Lagrange multiplier λ_0, which corresponds to the objective function vanishes. It may happen if the constraints of the problem are not regular at an extremal point.

*Dipartimento di Matematica per le Decisioni, University of Florence, Italy

†Departamento da Matemática & CEOC - Centro de Estudos em Optimização e Controlo, University of Aveiro, Portugal

To a given extremal point $\hat{x}$ there may correspond different sets $(\lambda_0, \overline{\lambda})$ of Lagrange multipliers, or, equivalently different extremals $(\hat{x}, \lambda_0, \overline{\lambda})$. Here Lagrange multiplier $\lambda_0 \in \mathbb{R}$ corresponds to objective function, $\overline{\lambda}$ is a vector of Lagrange multipliers corresponding to restrictions. If for some (each) of these extremals the multiplier λ_0 vanishes, the extremal point $\hat{x}$ is called (strictly) abnormal.

The traditional approach to optimization problems considers the fact that the objective function does not appear in optimal conditions as a "pathology". Common opinion was that in the cases of abnormal extremals it is not possible to deduce the optimality from the classical optimality conditions, and that further analysis is meaningless. That's why regularity conditions or some so called constraints qualification conditions that guarantee the normality of extremum are included into optimality conditions.

The problem of abnormal extremum is well known also in the Calculus of Variations. For the Lagrange problem:

$$\int_0^T \varphi(x(t), u(t))dt \longrightarrow min, \quad \dot{x} = f(x, u), \quad x(0) = x^0, x(T) = x^1,$$

the Lagrange multiplier λ_0, that corresponds to the integrand $\varphi(x, u)$, may vanish. It has been observed since long ago, that in this case the extremal trajectory $x(\cdot)$ may happen to be rigid. This means that $x(t)$ can not be locally varied without violating either the dynamic constraints or boundary conditions. At this point traditional approach based on small variations of $x(\cdot)$ seemingly fails. This phenomenon has been characterized by L.C.Young as "sad facts of life" [9]; the study of these "facts" has not been evolving for a long time.

Meanwhile the necessity to work with abnormal extrema emerged in different areas of optimization and control, as well as in geometry. The interest to this field has been very much stimulated by a discovery of R.Montgomery. Working with sub-Riemannian length-minimization problems he constructed ([8]) an example in which the only minimizing sub-Riemannian geodesic is strictly abnormal.

In the publications [1, 2] the abnormal extremum for the problem of minimizing of sub-Riemannian geodesics and more generally the problem of abnormal extremum has been approached from the point of view of (abnormal) sufficient second-order conditions. It has been established that the rigidity is a common phenomenon in optimization problems; in the abnormal case sufficient second-order optimality conditions imply rigidity of abnormal extremal points. In [1] the rigidity phenomenon has been studied for a class of Lagrange optimization problem in Banach space or equivalently for the NLP problems with equality constraints:

$$f(x) \to \min,\ H(x) = 0,$$

while in [2] rigidity of abnormal sub-Riemannian geodesics has been addressed. Besides necessary condition for rigidity has been proven.

In the present paper we address a more general NLP problem with mixed equality and inequality constraints

$$f(x) \to \min,\ H(x) = 0,\ G(x) \leq 0. \tag{1}$$

We restrict ourselves to the smooth case. The main result is proving that abnormal sufficient second-order optimality conditions for this problem imply rigidity of the abnormal extremal point.

The work of the authors has been partially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal - in the scope of the Program of Pluriannual Funding of CEOC - Centro de Estudos em Optimização e Controlo - of the University of Aveiro.

2 Second-Order Optimal Conditions for NLP Problem with Mixed Constraints: Sufficient Conditions

Consider the problem:

$$f(x) \longrightarrow min, \tag{2}$$

$$H(x) = 0, \tag{3}$$

$$G(x) \leq 0. \tag{4}$$

Here $x \in \mathbb{R}^n,\ f : \mathbb{R}^n \to \mathbb{R};\ \ H : \mathbb{R}^n \to \mathbb{R}^s,\ H(x) = [h_1(x), h_2(x), ..., h_s(x)]^T$; $G : \mathbb{R}^n \to \mathbb{R}^t, G(x) = [g_1(x), g_2(x), ..., g_t(x)]^T$.

The first order necessary optimality conditions at some point $\hat{x} \in \mathbb{R}^n$ amount to the existence of Lagrange multipliers $\hat{\lambda}_0 \geq 0,\ \hat{\lambda}_i \in \mathbb{R}, i = \overline{1,s};\ \hat{\mu}_j \geq 0, j = \overline{1,t}$ such that:

$$\hat{\lambda}_0 \nabla f(\hat{x}) + \sum_{i=1}^{s} \hat{\lambda}_i \nabla h_i(\hat{x}) + \sum_{j=1}^{t} \hat{\mu}_j \nabla g_j(\hat{x}) = 0 \tag{5}$$

together with the *complementary slackness conditions*:

$$\hat{\mu}_j g_j(\hat{x}) = 0, \quad j = \overline{1,t}. \tag{6}$$

Let us reunite the inequality constraints which are active at $\hat{x}$ into a map $\hat{G} : \mathbb{R}^n \to \mathbb{R}^{t_A}$(we assume that there are t_A active[1] inequality constraints at $\hat{x}$). The *regularity condition*:

$$Im(H'(\hat{x}), \hat{G}'(\hat{x})) = \mathbb{R}^t \times \mathbb{R}^{t_A}$$

guarantees nonvanishing of the Lagrange multiplier $\hat{\lambda}_0$. We will study now the general case, when $\hat{x}$ does not necessary satisfies the regularity condition and we assume that $\hat{\lambda}_0$ may vanish in (5). Our main result is the following theorem.

Theorem 1 (Sufficient conditions of normal and abnormal minimum) *Consider the problem (2) - (4) where the functions $f,\ h_i,\ g_j$ are twice differentiable in $\mathbb{R}^n$. Suppose that there exists some point $\hat{x} \in \mathbb{R}^n$ and some vector $(\hat{\lambda}_0, \hat{\lambda}, \hat{\mu}) \in \mathbb{R}^{s+t+1}, (\hat{\lambda}_0, \hat{\lambda}, \hat{\mu}) \neq 0,\ \hat{\lambda}_0 \geq 0,\ \hat{\mu} \geq 0;\ \hat{\lambda} = (\hat{\lambda}_i, i \in \mathbb{R}^s), \hat{\mu} = (\hat{\mu}_j, j \in \mathbb{R}^t)$ such that:*

i) $H(\hat{x}) = 0, \quad G(\hat{x}) \leq 0$;

ii) conditions (5) and the complementary slackness conditions (6) are satisfied, and in addition, for each $j = \overline{1,t}$ such that $g_j(\hat{x}) = 0$ the correspondent Lagrange multiplier $\hat{\mu}_j$ is positive: $\hat{\mu}_j > 0$ (modified complementary slackness conditions);

[1]If for some $j \in 1, 2, \ldots, t$ it is satisfied: $g_j(\hat{x}) = 0$ then the correspondent inequality constraint of problem (2)-(4) is called *active* at the point $\hat{x}$; otherwise the constraint is *passive* at $\hat{x}$.

iii) Hessian with respect to x of the Lagrange function:

$$\mathcal{L}(x, \lambda_0, \lambda, \mu) = \lambda_0 f(x) + \sum_{i=1}^{s} \lambda_i h_i(x) + \sum_{j=1}^{t} \mu_j g_j(x) \tag{7}$$

is positive definite over the subspace $M = \{\xi : H'(\hat{x})\xi = 0, \hat{G}'(\hat{x})\xi = 0\}$ *in* $(\hat{x}, \hat{\lambda}_0, \hat{\lambda}, \hat{\mu})$*, i.e., there exists* $\alpha > 0$ *such that*

$$\mathcal{L}''_{xx}(\hat{x}, \hat{\lambda}_0, \hat{\lambda}, \hat{\mu})(\xi, \xi) \geq \ 2\alpha \parallel \xi \parallel^2, \ \forall \xi \in M. \tag{8}$$

Then $\hat{x}$ supplies a strict local minimum to the problem (2) - (4).
If, in addition, $\hat{\lambda}_0 = 0$, then this minimum is rigid , i.e. the point $\hat{x}$ is isolated in the admissible set determined by the constraints: $S = \{x \in \mathbb{R}^n : H(x) = 0, G(x) \leq 0\}$.

Remark. The novelty of this result is in establishing the rigidity of abnormal extremal point under second order sufficient conditions of abnormal optimality. As far as we know it has not been pointed out before for the case of mixed constraints (compare with the corresponding condition in [4]).

Proof.

Without loss of generality we may assume: $\hat{x} = 0$ and that the inequality constraints are all *active* at $\hat{x} : g_j(\hat{x}) = 0, \ \forall j = \overline{1, t}$; then as we have assumed above $\hat{\mu}_j > 0$ for each $j = \overline{1, t}$.[2] Then the normal case of the Theorem 1 is a classical result, see [3].

Let us study the abnormal case when $\hat{\lambda}_0 = 0$. To establish the rigidity (isolatedness) of the point $\hat{x} = 0$ it will suffice to prove that for each x from some neighborhood of 0 in $\mathbb{R}^n$, $x \neq 0$, there holds:

$$\max\{\parallel H(x) \parallel, \ g_1(x), ..., g_t(x)\} > 0.$$

Here $\parallel H(x) \parallel = \max\limits_{i=\overline{1,s}} \mid h_i(x) \mid$.
We will prove moreover that there exists $\sigma > 0$ such that:

$$\max\{\parallel H(x) \parallel, g_1(x), ..., g_t(x)\} \geq \sigma \parallel x \parallel^2, \ \forall x \in S. \tag{9}$$

[2] Our treatment remains valid for the case where there are no inequality constraints.

The following Proposition provides an inequality from which (9) can be derived.

Proposition 1 *Suppose that there exists $\delta > 0$ such that for each $x \in S$:*

$$\mathcal{L}(x) + \max\{\| H(x) \|, g_1(x), ..., g_t(x)\} \geq \delta \| x \|^2, \tag{10}$$

where $\mathcal{L}(x) = \mathcal{L}(x, 0, \hat{\lambda}, \hat{\mu})$. Then the inequality (9) is valid for some $\sigma > 0$.

Proof of the Proposition 1. It is easy to conclude that (10) implies (9). Indeed, since

$$\max(\| \hat{\lambda} \|, \max_{j=\overline{1,t}} \hat{\mu}_j) = m > 0,$$

then we get:

$$\mathcal{L}(x) = \hat{\lambda} H(x) + \sum_{j=1}^{t} \hat{\mu}_j g_j(x) \leq |\hat{\lambda} H(x)| + \sum_{j=1}^{t} \hat{\mu}_j g_j(x)$$

$$\leq m \| H(x) \| + mt \max_{j=\overline{1,t}} g_j(x) \leq (m + tm) \max\{\| H(x) \|, g_1(x), ..., g_t(x)\}.$$

Substituting this estimate in place of $\mathcal{L}(x)$ in (10), we obtain:

$$(1 + m + tm) \max\{\| H(x) \|, g_1(x), ..., g_t(x)\} \geq \delta \| x \|^2,$$

which implies (9) with $\sigma = \frac{\delta}{1+m+tm}$. □

To establish (10) we will estimate separately $\mathcal{L}(x)$, $\| H(x) \|$ and $\max\{g_1(x), ..., g_t(x)\}$. For the Lagrange function $\mathcal{L}(x)$ we have: $\mathcal{L}(0) = 0$ and $\mathcal{L}'_x(0) = 0$. Hence the Taylor expansion of this function in the neighborhood of the point $\hat{x} = 0$ starts with the quadratic term:

$$\mathcal{L}(x) = \frac{1}{2} \mathcal{L}''_{xx}(0)(x, x) + \bar{o}(1) \| x \|^2 . \tag{11}$$

Now we will split the space $\mathbb{R}^n$ into three subspaces. First, consider $KerH'|_0$ and its complement Z in $\mathbb{R}^n$: $\mathbb{R}^n = KerH'|_0 \oplus Z$. Let us coordinatize Z by x_1. Next we split $KerH'|_0$ into subspace $M = KerH'|_0 \bigcap KerG'|_0$ and its complement $Y : KerH'|_0 = (KerH'|_0 \bigcap KerG'|_0) \oplus Y$. We will coordinatize M by x_{00} and Y by x_{01}. Finally, each $x \in \mathbb{R}^n$ admits representation:

$$x = x_{00} + x_{01} + x_1, \tag{12}$$

where x_{00}, x_{01}, x_1 are the projections of x on the correspondent subspaces M, Y, Z: $x_{00} \in M, x_{01} \in Y, x_1 \in Z$.

It is evident that the condition (8) means the positive definiteness of $\mathcal{L}''_{xx}|_0$ restricted onto the space M:

$$\mathcal{L}''_{xx}(0)(x_{00}, x_{00}) \geq 2\alpha \parallel x_{00} \parallel^2, \tag{13}$$

for some $\alpha > 0$.

Lemma 1 *For some $a > 0$, $\beta > 0$*

$$\mathcal{L}(x) \geq \max(0,\ a \parallel x_{00} \parallel^2 - \beta(\parallel x_{01} \parallel + \parallel x_1 \parallel) \parallel x \parallel). \tag{14}$$

Proof of Lemma 1. Let $\hat{\lambda} \in \mathbb{R}^s$, $\hat{\mu} \in \mathbb{R}^t > 0$ be the Lagrange multipliers that satisfy the conditions of the Theorem; $x \in S$. From (11) we have:

$$\begin{aligned} (\hat{\lambda}, \hat{\mu}) \begin{bmatrix} H(x) \\ G(x) \end{bmatrix} &= \mathcal{L}(x) = \mathcal{L}(x) - \mathcal{L}(0) \\ &= \frac{1}{2}\mathcal{L}''_{xx}(0)(x, x) + \bar{o}(1)(\parallel x_{00} \parallel^2 + \parallel x_{01} \parallel^2 + \parallel x_1 \parallel^2) \end{aligned} \tag{15}$$

as $\parallel x \parallel \longrightarrow 0$.

Since $\mathcal{L}''_{xx}(0)(x, x)$ is quadratic, we obtain:

$$\begin{aligned} \mid \mathcal{L}''_{xx}(0)(x, x) - \mathcal{L}''_{xx}(0)(x_{00}, x_{00}) \mid &= \underline{O}(1) \parallel x \parallel \parallel x - x_{00} \parallel \\ &= \underline{O}(1) \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel). \end{aligned} \tag{16}$$

From (15), (16) we get for x in some neighborhood V of the origin in $\mathbb{R}^n$ and for some $\beta > 0$:

$$\begin{aligned} (\hat{\lambda}, \hat{\mu}) \begin{bmatrix} H(x) \\ G(x) \end{bmatrix} &= \frac{1}{2}\mathcal{L}''_{xx}(0)(x_{00}, x_{00}) + \bar{o}(1)(\parallel x_{00} \parallel^2 + \parallel x_{01} \parallel^2 + \parallel x_1 \parallel^2) \\ &+ \underline{O}(1) \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel) \geq \frac{1}{2}\mathcal{L}''_{xx}(0)(x_{00}, x_{00}) \\ &- \alpha(x) \parallel x_{00} \parallel^2 - \beta \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel), \end{aligned} \tag{17}$$

where $\alpha(x) \to 0$ as $\| x \| \to 0$.

Without loss of generality we may assume that for $x \in V$ there holds: $\alpha(x) \leq \frac{\alpha}{2}$ where α is introduced in (13). Then from (8) and (17) we arrive to:

$$(\hat{\lambda}, \hat{\mu}) \begin{bmatrix} H(x) \\ G(x) \end{bmatrix} \geq \frac{\alpha}{2} \| x_{00} \|^2 - \beta \| x \| (\| x_{01} \| + \| x_1 \|),$$

which implies (14) for $a = \frac{\alpha}{2}$. □

Coming back to (10), we estimate now $\| H(x) \|$.

First let us show that for some $c > 0$, $A > 0$ and the splitting $x = x_1 + x_{01} + x_{00}$ introduced in (12) the estimate:

$$\| H(x) \| \geq \max\{0, c \| x_1 \| - A(\| x_{01} \|^2 + \| x_{00} \|^2)\}, \tag{18}$$

is valid. Indeed, as long as $H'|_0$ maps Z injectively, then for some $c > 0$ we have: $\| H'|_0 x_1 \| \geq c' \| x_1 \|$. From the Taylor formula for $H(x)$ we get:

$$H(x) = H(0) + H'|_0(x) + \overline{O}(\| x_{00} \|^2 + \| x_{01} \|^2 + \| x_1 \|^2)$$
$$= H'|_0(x_1) + A(x)(\| x_{00} \|^2 + \| x_{01} \|^2 + \| x_1 \|^2)$$

where $A(x)$ is bounded in V. Then

$$\| H(x) \| \geq c \| x_1 \| - A(\| x_{00} \|^2 + \| x_{01} \|^2),$$

for some $A > 0$ and some $0 < c \leq c'$.

Finally, we will find an estimate for $\max\{g_1(x), g_2(x), ..., g_t(x)\}$ in (10), proving the following Lemma.

Lemma 2 *For some $b > 0, \nu > 0, \omega > 0$ there holds:*

$$\max\{g_1(x), g_2(x), ..., g_t(x)\} \geq \max\{0, b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2\}. \tag{19}$$

Proof of Lemma 2. Let us multiply the vanishing vector

$$0 = \sum_{i=1}^{s} \hat{\lambda}_i \nabla h_i|_0 + \sum_{j=1}^{t} \hat{\mu}_j \nabla g_j|_0$$

by an arbitrary $\xi \in Y$ where Y is a complement subspace to $M = KerH'(0) \cap KerG'(0)$ in $KerH'(0)$. Evidently $\nabla h_i|_0\xi = 0$ for each $i = \overline{1,s}$, and hence we obtain:

$$\sum_{j=1}^{t} \hat{\mu}_j \nabla g_j|_0\xi = 0. \tag{20}$$

It is evident also that $\forall \xi \in Y \ \exists j \in 1,2,\dots,t: \ \nabla g_j|_0\xi \neq 0$. Since $\hat{\mu}_j > 0, j = \overline{1,t}$ (condition *ii*) of the Theorem), then some of the values $\nabla g_j|_0\xi$ in the left-hand side of (20) are positive and some are negative. It means that for each ξ one can subdivide the index set $J = \{1,2,\cdots,t\}$ into three subsets: $J = J^+ \cap J^0 \cap J^-$ such that:

$$J^+ = \{j \in J : \nabla g_j|_0\xi > 0\}, J^- = \{j \in J : \nabla g_j|_0\xi < 0\},$$

$$J^0 = \{j \in J : \nabla g_j|_0\xi = 0\}.$$

Then in (20) we have:

$$\sum_{j\in J^+} \hat{\mu}_j \nabla g_j|_0\xi + \sum_{j\in J^-} \hat{\mu}_j \nabla g_j|_0\xi = 0. \tag{21}$$

Let us take $\| G'(0)\xi \| = \max_{j\in J}\{|\nabla g_j|_0\xi|\}$. It is evident that the maximum can be achieved for some index j from either J^+ or J^- and for each $j \in J^+ \cup J^-$ there exists $b_j > 0$ such that:

$$| \nabla g_j|_0\xi | \geq b_j \| \xi \|, \tag{22}$$

as ∇G is injective in Y.
a) Suppose first that for some $j_m \in J^+ : \ \nabla g_{j_m}|_0\xi = \max_{j\in J}\{|\nabla g_j|_0\xi|\}$. In this case we conclude:

$$\nabla g_{j_m}|_0\xi = \max_{j\in J}\{|\nabla g_j|_0\xi|\} = \max_{j\in J}\{\nabla g_j|_0\xi\} \geq b_{j_m} \| \xi \| .$$

Then $\max_{j\in J}\{\nabla g_j|_0\xi\} \geq b \| \xi \|$, for $b = b_{j_m}$.

b) Suppose now that $\max_{j\in J}\{\nabla g_j|_0\xi\} = | \nabla g_{j_k}|_0\xi |$, where $j_k \in J^-$. From (20) written in the form:

$$\sum_{j\in J^+} \hat{\mu}_j \nabla g_j|_0\xi + \sum_{j\in J^-\setminus\{j_k\}} \hat{\mu}_j \nabla g_j|_0\xi + \hat{\mu}_{j_k} \nabla g_{j_k}|_0\xi = 0,$$

we get:

$$0 < -\hat{\mu}_{j_k} \nabla g_{j_k}|_0 \xi = \sum_{j \in J^+} \hat{\mu}_j \nabla g_j|_0 \xi + \sum_{j \in J^- \setminus \{j_k\}} \hat{\mu}_j \nabla g_j|_0 \xi.$$

Taking into account (22) we obtain the following estimate:

$$\hat{\mu}_{j_k} b_{jk} \parallel \xi \parallel \leq -\hat{\mu}_{j_k} \nabla g_{j_k}|_0 \xi = \sum_{j \in J^+} \hat{\mu}_j \nabla g_j|_0 \xi + \sum_{j \in J^- \setminus \{j_k\}} \hat{\mu}_j \nabla g_j|_0 \xi$$

$$\leq \sum_{j \in J^+} \hat{\mu}_j \nabla g_j|_0 \xi \leq \max_{j \in J^+} \{\nabla g_j|_0 \xi\} \sum_{j \in J^+} \hat{\mu}_j.$$

Therefore we obtain an inequality:

$$\max_{j \in J^+} \{\nabla g_j|_0 \xi\} \geq \frac{\hat{\mu}_{j_k} b_{jk}}{\sum_{j \in J^+} \hat{\mu}_j} \parallel \xi \parallel .$$

Assuming $b = \frac{\hat{\mu}_{j_k} b_{jk}}{\sum_{j \in J^+} \hat{\mu}_j} > 0$, we conclude that for each $\xi \in Y$:

$$\max_{j \in J} \{\nabla g_j|_0 \xi\} \geq b \parallel \xi \parallel . \tag{23}$$

We shall find now an upper estimate for the maximum of the functions $g_j(x), j \in J$. We invoke Taylor expansions of the functions $g_j(x), j \in J$ on the subspaces $M = KerH'(0) \cap KerG'(0)$, Y and Z. Recall that $g_j(0) = 0, \forall j \in J$ and that $x = x_{00} + x_{01} + x_1,\ x_{00} \in M,\ x_{01} \in Y,\ x_1 \in Z$ (representation (12)). Then

$$\begin{aligned} g_j(x_{00} + x_{01} + x_1) = \nabla g_j|_0(x_{00} + x_{01} + x_1) \\ + \underline{O}(1)(\parallel x_{00} \parallel^2 + \parallel x_{01} \parallel^2 + \parallel x_1 \parallel^2). \end{aligned} \tag{24}$$

It is evident that for all $j \in J$ and for all $x \in V$: $\nabla g_j|_0 x_{00} = 0$, and $| \nabla g_j|_0 x_1 | \leq \nu \parallel x_1 \parallel$ for some positive ν independent on x and j. Besides, we know from (23) that $\max_{j \in J} \{\nabla g_j|_0 x_{01}\} \geq b \parallel x_{01} \parallel$ for all $x \in V$. Then from (24) we conclude:

$$\max_{j \in J} \{g_j(x)\} = \max_{j \in J} \{g_j(x_{00} + x_{01} + x_1)\} \geq b \parallel x_{01} \parallel - \nu \parallel x_1 \parallel - \omega \parallel x_{00} \parallel^2,$$

where $\omega > 0$ is independent on x. Lemma 2 is proved. □

Now we are ready to estimate the expression in left-hand side of (10).
From (18), (19) we obtain:

$$\begin{gathered}\max\{\| H(x) \|,\ g_1(x), ...,\ g_t(x)\} \geq \max\{0, \\ c \| x_1 \| - A(\| x_{01} \|^2 + \| x_{00} \|^2), b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2\}.\end{gathered} \tag{25}$$

Summing the last inequality with (14) we obtain:

$$\begin{aligned}&\mathcal{L}(x) + \max\{\| H(x) \|,\ g_1(x), ..., g_t(x)\} \geq \max\{0, c \| x_1 \| \\ &- A(\| x_{01} \|^2 + \| x_{00} \|^2),\ b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2\} \\ &+ \max\{0,\ a \| x_{00} \|^2 - \beta \| x \| (\| x_{01} \| + \| x_1 \|)\}.\end{aligned} \tag{26}$$

To prove (10) we shall estimate the expression at the right-hand side of (26). Let $x \neq 0$. As x belongs to some neighborhood V of 0, we can assume:

$$\| x_1 \| \leq \varepsilon, \| x_{01} \| \leq \varepsilon, \| x_{00} \| \leq \varepsilon, \tag{27}$$

where the value of ε will be specified below. Let us obtain the estimates for

$$c \| x_1 \| - A(\| x_{01} \|^2 + \| x_{00} \|^2), \tag{28}$$

$$b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2, \tag{29}$$

$$a \| x_{00} \|^2 - \beta \| x \| (\| x_{01} \| + \| x_1 \|), \tag{30}$$

which are involved in the right side of (26).
Suppose first that

$$\frac{c}{2} \| x_1 \| \geq A(\| x_{01} \|^2 + \| x_{00} \|^2). \tag{31}$$

Then we have:

$$c \| x_1 \| - A(\| x_{01} \|^2 + \| x_{00} \|^2) \geq \frac{c}{2} \| x_1 \| = \frac{c}{4} \| x_1 \| + \frac{c}{4} \| x_1 \|$$

$$\geq \frac{c}{4} \| x_1 \| + \frac{A}{2}(\| x_{01} \|^2 + \| x_{00} \|^2) = \frac{c}{4 \| x_1 \|} \| x_1 \|^2$$

Cadernos de Matemática, Universidade de Aveiro, Portugal – Série de Investigação – 10 de Fevereiro de 2012

$$+\frac{A}{2}(\| x_{01} \|^2 + \| x_{00} \|^2) \geq \frac{c}{4\varepsilon} \| x_1 \|^2 + \frac{A}{2} \| x_{01} \|^2 + \frac{A}{2} \| x_{00} \|^2 .$$

If

$$\varepsilon \leq \frac{c}{2A} \tag{32}$$

then $\frac{c}{4\varepsilon} \geq \frac{A}{2}$ and (28) can be estimated as follows:

$$c \| x_1 \| - A(\| x_{01} \|^2 + \| x_{00} \|^2) \geq \frac{A}{2}(\| x_{00} \|^2 + \| x_{01} \|^2 + \| x_1 \|^2).$$

Suppose now that (31) is not satisfied, i.e.:

$$\| x_{01} \|^2 + \| x_{00} \|^2 > \frac{c}{2A} \| x_1 \| . \tag{33}$$

Then

$$b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2 = b \| x_{01} \| - \nu \frac{2A}{c}\frac{c}{2A} \| x_1 \| - \omega \| x_{00} \|^2$$

$$> b \| x_{01} \| - \frac{2A\nu}{c}(\| x_{01} \|^2 + \| x_{00} \|^2) - \omega \| x_{00} \|^2$$

$$= b \| x_{01} \| - \frac{2A\nu}{c} \| x_{01} \|^2 - \left(\frac{2A\nu}{c} + \omega\right) \| x_{00} \|^2$$

$$\geq (b - \frac{2A\nu\varepsilon}{c}) \| x_{01} \| - \gamma \| x_{00} \|^2,$$

where $\gamma = \frac{2A\nu}{c} + \omega$. Therefore, we obtain the estimate:

$$b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2 > (\frac{b}{2} + \frac{b}{2} - \frac{2A\varepsilon\nu}{c}) \| x_{01} \| - \gamma \| x_{00} \|^2 . \tag{34}$$

If

$$\varepsilon \leq \frac{cb}{4A\nu}, \tag{35}$$

then $\frac{b}{2} \geq \frac{2A\varepsilon\nu}{c}$ and we derive from (34):

$$b \| x_{01} \| - \nu \| x_1 \| - \omega \| x_{00} \|^2 > \frac{b}{2} \| x_{01} \| - \gamma \| x_{00} \|^2$$

$$= \frac{b}{4} \| x_{01} \| + \frac{b}{4} \| x_{01} \| - \gamma \| x_{00} \|^2 .$$

If, in addition to (33) there also holds

$$\frac{b}{4} \parallel x_{01} \parallel \geq \gamma \parallel x_{00} \parallel^2, \tag{36}$$

we will have:

$$b \parallel x_{01} \parallel -\nu \parallel x_1 \parallel -\omega \parallel x_{00} \parallel^2 \geq \frac{b}{4} \parallel x_{01} \parallel$$

$$= \frac{b}{8} \parallel x_{01} \parallel + \frac{b}{8} \parallel x_{01} \parallel \geq \frac{b}{8\varepsilon} \parallel x_{01} \parallel^2 + \frac{\gamma}{2} \parallel x_{00} \parallel^2 .$$

For $\frac{b}{8\varepsilon} \geq \frac{\gamma}{2}$, or equivalently for

$$\varepsilon \leq \frac{b}{4\gamma}, \tag{37}$$

we conclude:

$$b \parallel x_{01} \parallel -\nu \parallel x_1 \parallel -\omega \parallel x_{00} \parallel \geq \frac{\gamma}{2} \parallel x_{01} \parallel^2 + \frac{\gamma}{2} \parallel x_{00} \parallel^2$$

$$= \frac{\gamma}{4}(\parallel x_{01} \parallel^2 + \parallel x_{00} \parallel^2) + \frac{\gamma}{4}(\parallel x_{01} \parallel^2 + \parallel x_{00} \parallel^2).$$

Taking in account inequality (33) we continue the estimation:

$$b \parallel x_{01} \parallel -\nu \parallel x_1 \parallel -\omega \parallel x_{00} \parallel^2 > \frac{\gamma}{4}(\parallel x_{01} \parallel^2 + \parallel x_{00} \parallel^2) + \frac{\gamma}{4}\frac{c}{2A} \parallel x_1 \parallel$$

$$> \frac{\gamma}{4}(\parallel x_{01} \parallel^2 + \parallel x_{00} \parallel^2) + \frac{c\gamma}{8A\varepsilon} \parallel x_1 \parallel^2,$$

which implies an estimation for (29):

$$b \parallel x_{01} \parallel -\nu \parallel x_1 \parallel -\omega \parallel x_{00} \parallel^2 > \frac{\gamma}{4}(\parallel x_{01} \parallel^2 + \parallel x_{00} \parallel^2 + \parallel x_1 \parallel^2),$$

if

$$\varepsilon \leq \frac{c}{2A}. \tag{38}$$

Finally, if the inequality (33) holds, while (36) does not hold, i.e.:

$$\parallel x_{01} \parallel < \frac{4\gamma}{b} \parallel x_{00} \parallel^2, \tag{39}$$

then from (33) we obtain:

$$\frac{c}{2} \parallel x_1 \parallel < A \parallel x_{01} \parallel^2 + A \parallel x_{00} \parallel^2$$

$$\leq A(\frac{4\gamma}{b})^2 \parallel x_{00} \parallel^4 + A \parallel x_{00} \parallel^2 < A\frac{16\gamma^2\varepsilon^2}{b^2} \parallel x_{00} \parallel^2 + A \parallel x_{00} \parallel^2 .$$

Recall that by (37) $16\gamma^2\varepsilon^2 \leq b^2$ and hence the previous inequality results in

$$\parallel x_1 \parallel < \frac{4A}{c} \parallel x_{00} \parallel^2 . \tag{40}$$

For $\theta = \max\{\frac{4\gamma}{b},\ \frac{4A}{c}\}$ we derive from (39), (40):

$$\parallel x_{01} \parallel < \theta \parallel x_{00} \parallel^2, \parallel x_1 \parallel < \theta \parallel x_{00} \parallel^2 . \tag{41}$$

Taking in account (27), we can estimate (30) as follows:

$$a \parallel x_{00} \parallel^2 - \beta \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel) > a \parallel x_{00} \parallel^2 - \beta \parallel x \parallel 2\theta \parallel x_{00} \parallel^2 =$$

$$= (a - 2\beta\theta \parallel x \parallel) \parallel x_{00} \parallel^2 \geq (a - 6\beta\theta\varepsilon) \parallel x_{00} \parallel^2 .$$

Assuming: $\frac{a}{2} > 6\beta\theta\varepsilon$, or

$$\varepsilon < \frac{a}{12\beta\theta}, \tag{42}$$

we obtain

$$a \parallel x_{00} \parallel^2 - \beta \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel) \geq \frac{a}{2} \parallel x_{00} \parallel^2 .$$

Using the inequalities (41), (27) we continue:

$$a \parallel x_{00} \parallel^2 - \beta \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel) > 3\frac{a}{6} \parallel x_{00} \parallel^2$$

$$> \frac{a}{6}(\parallel x_{00} \parallel^2 + \frac{1}{\theta} \parallel x_1 \parallel + \frac{1}{\theta} \parallel x_{01} \parallel)$$

$$\geq \frac{a}{6}(\parallel x_{00} \parallel^2 + \frac{1}{\theta\varepsilon} \parallel x_1 \parallel^2 + \frac{1}{\theta\varepsilon} \parallel x_{01} \parallel^2)$$

$$= \frac{a}{6} \parallel x_{00} \parallel^2 + \frac{a}{6\theta\varepsilon} \parallel x_{01} \parallel^2 + \frac{a}{6\theta\varepsilon} \parallel x_{01} \parallel^2,$$

that gives an estimation for (30):

$$a \parallel x_{00} \parallel^2 -\beta \parallel x \parallel (\parallel x_{01} \parallel + \parallel x_1 \parallel) > K(\parallel x_{00} \parallel^2 + \parallel x_1 \parallel^2 + \parallel x_{01} \parallel^2),$$

where

$$K = \min\{\frac{a}{6},\ \frac{a}{6\theta\varepsilon}\}.$$

Therefore, choosing ε that satisfies all the inequalities (32), (35), (37), (42) we have (10) proved with $\delta = \min\{\frac{A}{2}, \frac{\gamma}{4}, K\} > 0$ for each x which satisfies inequalities (27).

□

References

[1] Agrachev A.A., Sarychev A.V. *On abnormal Extremals for Lagrange Variational Problems*, Journal of Mathematical Systems, Estimation and Control, (1998) Vol.8 , p.87-118.

[2] Agrachev A.A., Sarychev A.V. *Abnormal Subriemanian Geodesics: Morse Index and Rigidity*, Ann.Inst.Henri Poincare–Analyse Non Lineaire, (1996) Vol.13, p.635-690.

[3] Alekseev V.M., Tikhomirov V.M., Fomin S.V. *Optimal Control*, Plenum Publications (1987)

[4] Arutiunov A.V. *Optimality Conditions: Abnormal and Degenerate Problems*, Kluwer Academic Publishers (2000)

[5] John F. *Extremum problems with inequalitiess as subsidary constraints*, In K.O. Friedrichs, O.E. Neugebauer, and J.J. Stoker, editors, Studies and Essays Presented to R. Courant on His 70th Birthday Interscience Publishers, New York (1948), p.187-204.

[6] Kuhn H.W., Tucker A.W. *Non-linear Programming*, Proc. of the Second Berkley Symp. on Math. Stat. and Prob.- Berkley, Calif., (1951), p.481-492.

[7] Luenberger D.G. *Introduction to Linear and Nonlinear Programming*, Reading Mass.- London, Addison-Wesley (1973) V.12.

[8] Montgomery R. *Abnormal Minimizers*, SIAM J. Control Optimization, (1994) Vol.32, p.1605-1620.

[9] Young L.C. *Lectures on the Calculus of Variations and Optimal Control Theory*, Chelsea Publishing Company (1980)