SECOND ORDER CONDITIONS OF OPTIMALITY FOR CONSTRAINED OPTIMIZATION PROBLEMS IN FINITE DIMENSIONAL SPACES

by

E. Polak
E. J. Messerli

Memorandum No. ERL-M224
22 September 1967

ELECTRONICS RESEARCH LABORATORY
College of Engineering
University of California, Berkeley
94720

Second Order Conditions of Optimality for Constrained Optimization Problems in Finite Dimensional Spaces

by

E. Polak and E. J. Messerli

ABSTRACT

First-order necessary conditions of optimality for many problems in optimal control, nonlinear programming, and the calculus of variations can be obtained by transcribing these problems into a simple canonical form. In finite dimensional space this form reads:

(1) Basic Problem: Let $f : E^n \to E^1$, $r : E^n \to E^m$ be continuously differentiable functions and Ω a given subset of E^n. Find $\hat{x} \in \Omega$ such that $r(\hat{x}) = 0$, and, for all $x \in \Omega$ satisfying $r(x) = 0$, $f(\hat{x}) \leq f(x)$.

Roughly, the most general necessary condition for the Basic Problem (1) is of the form:

(2) If $\hat{x}$ is an optimal solution to the Basic Problem (1), then there is a nonzero vector $\psi \in E^{m+1}$ with $\psi^o \leq 0$ such that

$$\langle \psi^o \nabla f(\hat{x}) + \sum_{i=1}^{m} \psi^i \nabla r^i(\hat{x}), \delta x \rangle \leq 0$$ for all δx in a convex cone which "approximates" the set Ω at the optimal solution $\hat{x}$.

This general necessary condition can be satisfied trivially in some cases, such as when the gradients $\nabla f(\hat{x})$, $\nabla r^1(\hat{x}), \ldots, \nabla r^m(\hat{x})$ are linearly dependent, leading to a need for auxiliary necessary conditions for these cases.

In this paper the special case of linear dependence caused by the gradient $\nabla f(\hat{x})$ being zero is investigated. A new condition, involving second order partial derivatives of the cost function $f(\cdot)$, and first order partial derivatives of the equality constraint function $r(\cdot)$, is obtained. This new condition holds for all perturbation vectors in the convex cone associated with the general (first order) necessary condition (2) - a set which is larger than the one usually considered in obtaining second order conditions.

The research reported herein was supported by the National Aeronautics and Space Administration under Grant NsG-354, Supplement 4.

The authors are with the Electronics Research Laboratory, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley.

Second Order Conditions of Optimality for Constrained Optimization Problems in Finite Dimensional Spaces

by

E. Polak and E. J. Messerli

Introduction: In the last few years it has been shown [1, 2] that problems of the calculus of variations, nonlinear programming and optimal control can be treated in a unified manner as far as necessary conditions of optimality are concerned. This was done by establishing that all these problems can be transcribed into a simple canonical form, for which necessary conditions were developed. Specialized necessary conditions of optimality for any particular problem then followed from the structure of the problem.

For finite dimensional problems the canonical form mentioned above reads as follows:

1. Basic Problem: Let $f : E^n \to E^1$, $r : E^n \to E^m$ be continuously differentiable functions, and let Ω be a subset of E^n. Find a vector $\hat{x} \in \Omega$ such that: (i) $r(\hat{x}) = 0$ and (ii) for every x in Ω with $r(x) = 0$, $f(x) \geq f(\hat{x})$.

Thus, for example, the usual Nonlinear Programming Problem, $\min\{f(x) | r(x) = 0,\ q(x) \leq 0\}$, where $f:E^n \to E^1$, $r:E^n \to E^m$ and $q:E^n \to E^k$ are continuously differentiable, is recognized to be the Basic Problem (1) with $\Omega = \{x | q(x) \leq 0\}$. Examples of the transcription of discrete optimal control problems to Basic Problem form may be found in [1].

Before giving the necessary condition for the Basic Problem (1) we require an "approximation" of the set Ω at a given point.

2. Definition: A convex cone $C(\hat{x}, \Omega)$ will be called a conical approximation to the constraint set Ω at $\hat{x}$ if for any collection $\{\delta x_1, \ldots, \delta x_k\}$ of linearly independent vectors in $C(\hat{x}, \Omega)$ there exists an $\varepsilon > 0$ (possibly depending on $\hat{x}$, $\delta x_1, \ldots, \delta x_k$), and a continuous map $\zeta(\cdot)$ from the convex hull (co) of $\{0, \delta x_1, \ldots, \delta x_k\}$ into $\Omega - \hat{x}$ such that $\zeta(\delta x) = \varepsilon\,\delta x + o(\varepsilon\,\delta x)$ where $\|o(\delta x)\| / \|\delta x\| \to 0$ as $\|\delta x\| \to 0$.

The most general necessary condition for the Basic Problem (1) is the following one.

3. Fundamental Theorem [1]: If $\hat{x}$ is a solution to the Basic Problem (1), and $C(\hat{x}, \Omega)$ is a conical approximation to Ω at $\hat{x}$, then there exists a nonzero vector $\psi = (\psi^o, \ldots, \psi^m)$ in E^{m+1} with $\psi^o \leq 0$ such that for every δx in the closure, $\overline{C(\hat{x}, \Omega)}$, of $C(\hat{x}, \Omega)$:

4. $$\langle \psi^o \nabla f(\hat{x}) + \sum_{i=1}^{m} \lambda^i \nabla r^i(\hat{x}),\ \delta x \rangle \leq 0$$

The research reported herein was supported by the National Aeronautics and Space Administration under Grant NsG-354, Supplement 4.

The authors are with the Electronics Research Laboratory, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley.

Note that the Fundamental Theorem (3) may become degenerate in two ways. The first occurs when ψ^o must be chosen to be zero, and hence no information about the cost function $f(\cdot)$ enters into the necessary condition (4). This most often occurs when there is only one $x \in \Omega$ satisfying $r(x) = 0$, and may be avoided by introducing a regularity condition, such as the Kuhn-Tucker constraint qualification, on $r(\cdot)$ and Ω. The Fundamental Theorem also becomes degenerate when the vectors $\nabla f(\hat{x})$, $\nabla r^1(\hat{x}), \dots, \nabla r^m(\hat{x})$ are linearly dependent since then one can always choose a $\psi \neq 0$ which satisfies $\psi^o \nabla f(\hat{x}) + \sum_{i=1}^{m} \psi^i \nabla r^i(\hat{x}) = 0$, and hence (4), without reference to the optimality of $\hat{x}$.

When a degeneracy in the first-order condition occurs, it is obviously desirable to have a second-order necessary condition. However, there are other cases when a second-order condition is also meaningful. Thus, suppose that in $C(\hat{x}, \Omega)$ there are "critical" vectors y which satisfy $\langle \nabla f(\hat{x}), y \rangle = 0$ and $\langle \nabla r^i(\hat{x}), y \rangle = 0 \quad i = 1, \dots, m$. Then, under suitable assumptions, one obtains for these vectors relations of the form:

5. $$y^T \frac{\partial^2 f}{\partial x^2} (\hat{x})\, y \geq 0$$

or

6. $$y^T \left(\frac{\partial^2 f}{\partial x^2} (\hat{x}) - \sum_{i=1}^{m} \lambda^i \frac{\partial^2 r^i}{\partial x^2} (\hat{x}) - \sum_{i=1}^{k} u^i \frac{\partial^2 q^i}{\partial x^2} (\hat{x}) \right) y \geq 0$$

(see, for example [3], [4]).

In this paper we consider a special case of degeneracy in the first-order condition (4), namely the case when $\nabla f(\hat{x}) = 0$, which causes the vectors $\nabla f(\hat{x})$, $\nabla r^1(\hat{x}), \nabla r^2(\hat{x}), \dots, \nabla r^m(\hat{x})$ to be linearly dependent. However, we shall not restrict ourselves to critical directions only as in [3], [4], and, instead, we shall obtain a condition similar to (4), but with $\delta x^T \frac{\partial^2 f}{\partial x^2} (\hat{x})\, \delta x$ playing the role of $\langle \nabla f(\hat{x}), \delta x \rangle$.

II. A Second-Order Condition: Let us assume that $\hat{x}$ is a solution to the Basic Problem (1) such that $\nabla f(\hat{x}) = 0$, and suppose that $f(\cdot)$ is twice continuously differentiable. Then to the Fundamental Theorem (3) we can add the following new second-order condition:

7. Theorem: If $\hat{x}$ is a solution to the Basic Problem (1) such that $\nabla f(\hat{x}) = 0$, and $C(\hat{x}, \Omega)$ is a conical approximation of Ω at the point $\hat{x}$, then the ray R,

8. $$R \triangleq \{ y \in E^{m+1} \mid y = \beta(-1, 0, 0, 0 \dots, 0) \quad \beta \geq 0 \}$$

has no points in the interior of the set L defined by:

9. $$L \triangleq \{ (y^o, y) \mid y^o = \delta x^T \frac{\partial^2 f}{\partial x^2} (\hat{x})\, \delta x, \ y = \frac{\partial r}{\partial x} (\hat{x})\, \delta x; \ \delta x \in C(\hat{x}, \Omega) \}$$

An equivalent statement of the Fundamental Theorem (3) is that the ray R given by (8) has no points in the interior of the set L_o defined by:

10. $$L_o \triangleq \{ (y^o, y) \mid y^o = \langle \nabla f(\hat{x}), \delta x \rangle, \ y = \frac{\partial r}{\partial x} (\hat{x})\, \delta x, \ \delta x \in C(\hat{x}, \Omega) \}.$$

Since L_o is also a convex cone, L_o and R must be separated, which is the essence of the statement of Theorem (3) in the original form given.

Since L defined in (9) is not in general convex it is natural to inquire if there is a curved surface which separates R and L - rather than a plane. In fact a parboloid of the form:

11. $$\tilde{g}_\lambda(y) = \lambda^o y^o + \sum_{i=1}^{m} \lambda^i (y^i)^2 \qquad \text{with} \qquad \lambda^o \leq 0$$

is the logical candidate, which, on substituting $y \in L$ gives the following quadratic in δx,

12. $$g_\lambda(\delta x) = \lambda^o \delta x^T \frac{\partial^2 f}{\partial x^2} \delta x + \sum_{i=1}^{m} \lambda^i \delta x^T (\nabla r^i(\hat{x})) \langle \nabla r^i(\hat{x})) \delta x .$$

We are thus led to the following consequence of Theorem (7).

13. Theorem: If $\hat{x}$ is a solution to the Basic Problem (1) such that $\nabla f(\hat{x}) = 0$, and $C(\hat{x}, \Omega)$ is a conical approximation of Ω at $\hat{x}$, then there is a nonzero vector $\lambda \in E^{m+1}$ with $\lambda^o \leq 0$ such that for every δx in $C(\hat{x}, \Omega)$,

14. $$\lambda^o \delta x^T \frac{\partial^2 f}{\partial x^2}(\hat{x}) \delta x + \sum_{i=1}^{m} \lambda^i \langle \nabla r^i(\hat{x}), \delta x \rangle^2 \leq 0$$

15. Moreover, if the ray R (8) is not a boundary ray of the set L (9), λ^o may be taken as -1.

16. Remark: The relation (14) can always be satisfied trivially if we allow $\lambda^o = 0$. While there are cases in which λ^o must be chosen to be zero, the qualification (15) allows a nontrivial statement for many problems.

Thus, let us again consider the Nonlinear Programming Problem,

17. $$\min \{f(x) \mid r(x) = 0, \; q(x) \leq 0\}$$

where $f: E^n \to E'$ is assumed twice continuously differentiable, and $r: E^n \to E^m$, $q: E^n \to E^k$ are continuously differentiable.

Define $I(x) = \{i \in \{1, \ldots, k\} \mid q^i(x) = 0\}$ and $IC(x) = \{y \mid \langle \nabla q^i(x), y \rangle < 0 \; i \in I(x)\}$.

We now obtain the following condition from Theorem (13).

18. Theorem: Suppose $\hat{x}$ is an optimal solution to the Nonlinear Programming Problem (17), $\nabla f(\hat{x}) = 0$, and there are positive constants ρ_1, ρ_2 such that $f(x) \geq f(\hat{x}) + \rho_1 ||x-\hat{x}||^2$ for all x satisfying $r(x) = 0$, $g(x) \leq 0$ and $||x-\hat{x}|| \leq \rho_2$. If in addition $IC(\hat{x}) \neq \phi$, and the Kuhn-Tucker Constraint Qualification [5] is satisfied, then there is a scalar λ such that

19. $$-y^T \frac{\partial^2 f}{\partial x^2}(\hat{x}) y + \lambda \sum_{i=1}^{m} \langle \nabla r^i(\hat{x}), y \rangle^2 \leq 0 \quad \text{for all } y \in \overline{IC(\hat{x})} .$$

Since the conditions of Theorem (18) may be difficult to verify, it is perhaps more meaningful to state the following local "sufficiency" condition.

20. Theorem: Suppose that $\hat{x}$ satisfies $r(\hat{x}) = 0$, $g(\hat{x}) \leq 0$, and $\nabla f(\hat{x}) = 0$. If there is a scalar λ such that

21. $$-y^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\, y + \lambda \sum_{i=1}^{m} \langle \nabla r^i(\hat{x}), \delta x \rangle^2 < 0 \text{ for all } \delta x \neq 0,$$

then there are positive constants ρ_1, ρ_2 such that $f(x) \geq f(\hat{x}) + \rho_1 ||x-\hat{x}||^2$ for all x satisfying $r(x) = 0$, $g(x) \leq 0$ and $||x-\hat{x}|| \leq \rho_2$.

III. Conclusions: We have shown in this paper that, when first-order necessary conditions of optimality fail because the gradient of the cost function at the optimal point is zero, it is possible to replace these first-order conditions with a new condition. This new condition, involving second-order partial derivatives of the cost function and first-order partial derivatives of the equality constraint function, holds for all perturbation vectors in a set which is larger than the one usually considered in obtaining second-order conditions.

References:

1. M. Canon, C. Cullum, and E. Polak, "Constrained Minimization Problems in Finite Dimensional Spaces," J. SIAM Control, Vol. 4, No. 3, 1966, pp. 528-547

2. N. Da Cunha and E. Polak, "Constrained Minimization under Vector-Valued Criteria in Linear Topological Spaces," in Mathematical Theory of Control, Academic Press, New York, 1967.

3. M. R. Hestenes, Calculus of Variations and Optimal Control Theory, Wiley, New York, 1966.

4. G. P. McCormick, "Second Order Conditions for Constrained Minima," J. SIAM App. Math., Vol. 15, No. 3, pp. 641-652, 1967.

5. H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," Proc. of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, California, 1951, pp. 481-492.

6. L. M. Graves, The Theory of Functions of a Real Variable, 2nd. ed., McGraw-Hill, New York, pp. 146-149, 1956.

Appendix: Proof of Theorems (7), (13), (18) and (20)

Theorem (7):

Let $\hat{x}$ be an optimal solution to the Basic Problem (1), with $\nabla f(\hat{x}) = 0$, and assume that Theorem (7) is false, i.e., that the ray R (8) has points in the interior of the set L (9). From this contrary assumption it follows that there are linearly independent vectors $\delta x_1, \ldots, \delta x_{m+1}$ in $C(\hat{x},\Omega)$, with corresponding map $\zeta(\cdot)$ defined as in (2), such that:

(i) $\zeta(\epsilon\,\delta x) \in \Omega - \hat{x}$ for all $\delta x \in co\{\delta x_1, \ldots, \delta x_{m+1}\}$ and $\epsilon \in [0,1]$.

(ii) The set $\Sigma = co\,\{\delta y_1, \ldots, \delta y_{m+1}\}$, where $\delta y_i = \frac{\partial r}{\partial x}(\hat{x})(\delta x_i)$ for $i = 1, \ldots, m+1$, is a simplex† in E^m, containing the origin in its interior.

(iii) $f(\hat{x} + \zeta(\epsilon\,\delta x)) < f(\hat{x})$ for all $\delta x \in co\,\{\delta x_1, \ldots, \delta x_{m+1}\}$ and $\epsilon \in (0,1]$.

The existence of $\delta x_1, \ldots, \delta x_{m+1}$ satisfying the above conditions requires some verification, which we now give. First, we observe that there is a vector $\widetilde{\delta x}$ in $C(\hat{x},\Omega)$ such that

$$\langle\, \widetilde{\delta x},\ \frac{\partial^2 f}{\partial x^2}(\hat{x})\,\widetilde{\delta x}\,\rangle = -1, \text{ and } \langle\, \nabla r^i(\hat{x}),\ \widetilde{\delta x}\,\rangle = 0$$

for $i = 1, \ldots, m$, and hence there is a ball $B(\widetilde{\delta x})$ about $\widetilde{\delta x}$ such that

$$\langle\, \delta x,\ \frac{\partial^2 f}{\partial x^2}(\hat{x})\,\delta x\,\rangle < -1/2$$

for all $\delta x \in B(\widetilde{\delta x})$.

Secondly, if vectors $\delta y_1', \ldots, \delta y_{m+1}'$ form a simplex, $co\,\{\delta y_1', \ldots, \delta y_{m+1}'\}$, in E^m, containing the origin in its interior, then there are vectors $\delta x_1', \ldots, \delta x_{m+1}'$ in $C(\hat{x}, \Omega)$ such that

$$\delta y_i' = \frac{\partial r}{\partial x}(\hat{x})\,\delta x_1' \text{ for } i = 1, \ldots, m+1.$$

We now define $\delta x_i = (1 - \lambda_i)\,\delta x_i' + \lambda_i\,\widetilde{\delta x}$ for $i = 1, \ldots, m+1$, where $\lambda_i \in [0,1)$ is chosen such that $\delta x_i \in B(\widetilde{\delta x})$. The vectors $\delta x_1, \ldots, \delta x_{m+1}$ are linearly independent since otherwise $0 \in co\,\{\delta x_1, \ldots, \delta x_{m+1}\}$, which is impossible since $co\,\{\delta x_1, \ldots, \delta x_{m+1}\} \subset B(\widetilde{\delta x})$. Thus, since

$$\langle\, \delta x,\ \frac{\partial^2 f}{\partial x^2}(\hat{x})\,\delta x\,\rangle < -1/2$$

for all $\delta x \in co\,\{\delta x_1, \ldots, \delta x_{m+1}\}$, and $\delta x_1, \ldots, \delta x_{m+1}$ are linearly independent, there is an $\epsilon_o > 0$ such that, with δx_i redefined as $\epsilon_o\,\delta x_i$ for $i = 1, \ldots, m+1$, conditions (i), (ii), and (iii) above are satisfied.

---

†A simplex in E^n is a convex polyhedron with n+1 vertices, which has a nonempty interior.

We now define the m x m matrix Y with i-th column $\delta y_i - \delta y_{m+1}$ for $i = 1, \ldots, m$ and the n x m matrix X with i-th column $\delta x_i - \delta x_{m+1}$ for $i = 1, \ldots, m$. Then Y is nonsingular since Σ is a simplex, and hence $XY^{-1}(\delta y - \delta y_{m+1}) + \delta x_{m+1}$ is a continuous map from Σ to $co\{\delta x_1, \ldots, \delta x_{m+1}\}$.

For $\alpha \in (0, 1]$ we now define the uniformly continuous map $G_\alpha : \Sigma \to E^m$ by:

20. $$G_\alpha(\delta y) = \delta y - 1/\alpha\ r(\hat{x} + \zeta\ (\alpha XY^{-1}(\delta y - \delta y_{m+1}) + \alpha \delta x_{m+1}))$$

which, on expansion, reduces to

21. $$G_\alpha(\delta y) = \frac{\tilde{o}(\alpha, \delta y)}{\alpha}$$

where $\|\tilde{o}(\alpha, \delta y)\|/\alpha \to 0$ as $\alpha \to 0$, uniformly for $\delta y \in \Sigma$. Thus, there is an $\alpha^* \in (0,1]$ such that $G_{\alpha^*}(\delta y) \in \Sigma$ for all $\delta y \in \Sigma$, and, therefore, from Brouwer's fixed point Theorem [6], there is a $\delta y^* \in \Sigma$ such that $G_{\alpha^*}(\delta y^*) = \delta y^*$.

From (20) we see that the point

$$x^* \triangleq \hat{x} + \zeta\ (\alpha^* XY^{-1}(\delta y^* - \delta y_{m+1}) + \alpha^* \delta x_{m+1})$$

satisfies $r(x^*) = 0$, and since by condition (i), $x^* \in \Omega$, and by condition (iii), $f(x^*) < f(\hat{x})$, $\hat{x}$ cannot be optimal, which is a contradiction.

Theorem (13): We shall only consider the case $\lambda^o = -1$, i.e. the ray R(8) is not a boundary ray of the set L(9). Thus, there is a closed ball B_ϵ, of radius ϵ, about (- 1, 0, . . . 0) such that $B_\epsilon \cap \bar{L} = \phi$.

Now consider the cone in E^2, $\{y \mid y^o = \delta x^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\ \delta x,$

$$y^1 = \sum_{i=1}^{m} \langle \nabla r^i(\hat{x}),\ \delta x \rangle^2,\ \delta x \in C(\hat{x})\Omega\ )\}\ .$$

Since $y^1 \geq 0$ for all δx this cone is separated from the ray R. If it is not separated strictly then there is a $\delta x \in C(\hat{x}, \Omega)$ such that

$$\delta x^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\ \delta x = -1 \text{ and } \sum_{i=1}^{m} \langle \nabla r^i(\hat{x}),\ \delta x \rangle^2 < \epsilon^2,$$

i.e. $(\langle \delta x^T, \frac{\partial^2 f}{\partial x^2}(\hat{x})\ \delta x \rangle, \langle \nabla r^1(\hat{x}),\ \delta x \rangle, \ldots, \langle \nabla r^m(\hat{x}),\ \delta x \rangle) \in B_\epsilon$,

which is a contradiction, and this completes the proof.

Theorem (18): The theorem follows providing R is not a boundary ray of the set $\bar{L}$, thus assume the contrary. It follows that there is a sequence $\{y_i\}$ of unit vectors in $IC(\hat{x})$, converging to $y_* \in \overline{IC(\hat{x})}$ with $\|y_*\| = 1$ such that

$$y_*^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\ y_* \leq 0 \text{ and } \langle \nabla r^i(\hat{x}),\ y_* \rangle = 0 \quad \text{for } i = 1, \ldots, m$$

By the Kuhn-Tucker constraint qualification, $r(\hat{x}+ty_* + o(t)) = 0$, and $q(\hat{x}+ty_* + o(t)) \leq 0$, $t \in [0, \bar{t}]$, for some $\bar{t} > 0$ and some continuous function $o(\cdot)$ satisfying

$$\lim_{t \to 0^+} \frac{\| o(t) \|}{t} = 0.$$

Since $f(x + ty + o(t))$ expands as:

$$f(\hat{x} + ty_* + o(t)) = f(\hat{x}) + t^2 y_*^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\, y_* + o'(t^2),$$

and $y_*^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\, y_* \leq 0$ with $y_* \neq 0$, ρ_1 and ρ_2 satisfying the conditions given in Theorem (18) cannot exist, which is a contradiction.

Theorem (20): Let x_i be a sequence (with $x_i \neq \hat{x}$ for any i) converging to $\hat{x}$ such that $r(x_i) = 0$ and $q(x_i) \leq$ for all i. Let

$$y_i = \left\{ \frac{x_i - \hat{x}}{\| x_i - \hat{x} \|} \right\},$$ $i \in K$, an index set, be any convergent subsequence of $\frac{x_i - \hat{x}}{\| x_i - \hat{x} \|}$, converging to y_*. Then $y_* \in \overline{IC(\hat{x})}$, $\langle \nabla r^i(\hat{x}), y_* \rangle = 0$ for $i = 1, \ldots, m$, and $\| y_* \| = 1$.

The set S of all such y_* generated in this way is compact, and hence by (21) there is a $\rho > 0$ such that

$$y_*^T \frac{\partial^2 f}{\partial x^2}(\hat{x})\, y_* \geq \rho = \rho \, \| y_* \|^2 > 0 \text{ for all } y_* \in S,$$ and this leads to Theorem (20).