Journal of Global Optimization **27:** 367–374, 2003.

On the Complexity of Equalizing Inequalities

HUBERTUS TH. JONGEN and OLIVER STEIN
Department of Mathematics – C, Aachen University, Aachen, Germany (e-mail: jongen@rwth-aachen.de)

(Received 14 October 2002; accepted in revised form 3 February 2003)

Abstract. We study two approaches to replace a finite mathematical programming problem with inequality constraints by a problem that contains only equality constraints. The first approach lifts the feasible set into a high-dimensional space by the introduction of quadratic slack variables. We show that then not only the number of critical points but also the topological complexity of the feasible set grow exponentially. On the other hand, the second approach bases on an interior point technique and lifts an approximation of the feasible set into a space with only one additional dimension. Here only Karush–Kuhn–Tucker points with respect to the positive and negative objective function in the original problem give rise to critical points of the smoothed problem, so that the number of critical points as well as the topological complexity can at most double.

Key words: Euler characteristic, interior point method, logarithmic smoothing, Morse theory, quadratic slack

1. Introduction

In this article we study optimization problems of the type

$$P: \qquad \text{minimize} \;\; f(x) \;\; \text{subject to } x \in M$$

where

$$M = \{x \in \mathbb{R}^n | h_i(x) = 0, i \in I, g_j(x) \geqslant 0, j \in J\}$$

and where the defining functions $f, h_i, g_j\colon \mathbb{R}^n \to \mathbb{R}$ are C^2-functions, $|I| < n, |J| < \infty$.

The presence of inequality constraints in P does not only complicate the boundary structure of its feasible set M, but it also constitutes a well-known challenge for the design and performance of numerical solution methods for P. The aim of this article is to study how P can be replaced by optimization problems *without* inequality constraints. In particular, we investigate an exact approach and an approximative approach via smoothing, both of which “lift” the problem into a higher-dimensional space. The complexity of the new problems is higher in the sense that the number of critical points increases in comparison to the original problem. We show that the number of critical points grows exponentially for the exact approach, whereas for the smoothing approach it doubles at most. Via Morse theory we also derive corresponding results on the topological complexity of the lifted sets.

The article is organized as follows. In Section 2 we recall the notions of a critical point, a Karush–Kuhn–Tucker point, a non-degenerate critical point, and of the linear and quadratic indices. Section 3 treats the exact approach of introducing quadratic slack variables. In Theorem 3.7 we show that then the number of critical points grows exponentially in the number of inactive inequality constraints, and Theorem 3.8 gives the relation between the occurring linear and active indices. By means of an example Theorem 3.10 stresses that also the topological complexity of the feasible set can grow exponentially. As opposed to these observations, Section 4 shows that the number of critical points of a logarithmically smoothed problem is bounded by twice the number of Karush–Kuhn–Tucker points with respect to the positive and negative objective function in the original problem (Theorem 4.4), and that also the topological complexity does not change drastically (Theorem 4.5).

2. Basic notions

At a point $\bar{x} \in M$ the *linear independence constraint qualification (LICQ)* is said to hold, if the vectors $Dh_i(\bar{x}),\ i \in I,\ Dg_j(\bar{x}),\ j \in J_0(\bar{x})$, are linearly independent. Here, DF stands for the row vector of partial derivatives of a real-valued function F, and $J_0(\bar{x}) = \{j \in J \mid g_j(\bar{x}) = 0\}$ denotes the set of active inequality constraints at $\bar{x}$. Generically, LICQ holds at every point in M and, in this case, M is a $(n - |I|)$-dimensional C^2-manifold with boundary (cf. [3]).

A point $\bar{x} \in M$ is called *critical point* for $f|_M$ if LICQ holds at $\bar{x}$ and if there exist real numbers (Lagrange multipliers) $\bar{\lambda}_i,\ i \in I,\ \bar{\mu}_j,\ j \in J_0(\bar{x})$, such that

$$Df(\bar{x}) \;=\; \sum_{i\in I} \bar{\lambda}_i\, Dh_i(\bar{x}) \;+\; \sum_{j\in J_0(\bar{x})} \bar{\mu}_j\, Dg_j(\bar{x})\ .$$

A critical point is called *Karush-Kuhn-Tucker point (KKT-point)* if $\bar{\mu}_j \geqslant 0$, $j \in J_0(\bar{x})$. For an optimization problem P without inequality constraints, i.e. $J = \emptyset$, the sets of critical points and of KKT-points obviously coincide.

A critical point is called *non-degenerate* if the following two conditions hold:

ND1: $\bar{\mu}_j \neq 0,\ j \in J_0(\bar{x})$,
ND2: $D^2L(\bar{x})|_{T_{\bar{x}}M}$ is non-singular .

The matrix D^2L stands for the Hessian of the Lagrange function

$$L(x) \;=\; f(x) \;-\; \sum_{i\in I} \bar{\lambda}_i\, h_i(x) \;-\; \sum_{j\in J_0(\bar{x})} \bar{\mu}_j\, g_j(x)\ , \tag{1}$$

and $T_{\bar{x}}M$ denotes the tangent space of M at $\bar{x}$,

$$T_{\bar{x}}M \;=\; \{\ \xi \in \mathbb{R}^n \mid Dh_i(\bar{x})\,\xi = 0,\ i \in I,\ Dg_j(\bar{x})\,\xi = 0,\ j \in J_0(\bar{x})\ \}\ .$$

Condition ND2 means that $V^\top D^2 L(\bar{x}) V$ is non-singular, where V is some matrix whose columns form a basis of the tangent space $T_{\bar{x}} M$. The number of negative (positive) multipliers $\bar{\mu}_j$ in ND1 is called the *linear index (LI)* (*linear co-index (LCI)*) of $\bar{x}$. The number of negative (positive) eigenvalues of $D^2 L(\bar{x})|_{T_{\bar{x}} M}$ in ND2 is the *quadratic index (QI)* (*quadratic co-index (QCI)*) of $\bar{x}$. In particular, a non-degenerate critical point is a KKT-point if and only if LI=0. In that case the quadratic index coincides with the so-called Morse-index. Note that a non-degenerate critical point is a local minimum (maximum) of P if and only if LI=QI=0 (LCI=QCI=0). Generically, all critical points of an optimization problem P are non-degenerate (cf. [3]).

3. An exact approach

The following approach to equalize inequality constraints by adding quadratic slack variables is well-known (see, e.g., [6] and [7]). Consider the problem

$$\widetilde{P}: \qquad \text{minimize } \; f(x) \; \text{ subject to } (x, z) \in \widetilde{M}$$

where

$$\widetilde{M} \; = \; \{ \, (x, z) \in \mathbb{R}^n \times \mathbb{R}^{|J|} \, | \; h_i(x) = 0, \; i \in I, \; g_j(x) - z_j^2 = 0, \; j \in J \, \} \, .$$

The following results are easily seen.

LEMMA 3.1. *If LICQ holds at $\bar{x} \in M$, then LICQ holds at $(\bar{x}, \bar{z}) \in \widetilde{M}$, where $\bar{z}_j = \pm\sqrt{2 g_j(\bar{x})}\,, \; j \in J$.*

COROLLARY 3.2. *If LICQ holds at all $x \in M$, then $\widetilde{M}$ is a C^2-manifold with* $\dim \widetilde{M} \; = \; \dim M$.

LEMMA 3.3. *M is compact if and only if $\widetilde{M}$ is compact.*

LEMMA 3.4. *Let LICQ hold at $\bar{x} \in M$ and put $\bar{z}_j = \pm\sqrt{2 g_j(\bar{x})}\,, \; j \in J$. Then the tangent space to $\widetilde{M}$ at $(\bar{x}, \bar{z})$ has the form*

$$T_{(\bar{x},\bar{z})} \widetilde{M} \; = \; \{ \, (\xi, \eta) \in T_{\bar{x}} M \times \mathbb{R}^{|J|} | \eta_j \in \mathbb{R} \, , \; j \in J_0(\bar{x}) \, , \qquad \eta_j = \tfrac{1}{\bar{z}_j} \, D g_j(\bar{x}) \, \xi \; , \; j \in J \setminus J_0(\bar{x}) \, \} \, .$$

From now on let the following assumption hold:

ASSUMPTION 3.5. *LICQ holds at each point $x \in M$.*

LEMMA 3.6. *Let $\bar{x} \in M$ be a critical point for $f|_M$ with Lagrange multipliers $\bar{\lambda}_i \, , \; i \in I, \, \bar{\mu}_j \, , \; j \in J_0(\bar{x})$. Then $(\bar{x}, \bar{z})$ with $\bar{z}_j = \pm\sqrt{2 g_j(\bar{x})}\,, \; j \in J$, is a critical* *point for* $f|_{\widetilde{M}}$ *(*and hence KKT-point*), i.e.*

$$\begin{pmatrix} D^\top f(\bar{x}) \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix} = \sum_{i\in I} \bar{\lambda}_i \begin{pmatrix} D^\top h_i(\bar{x}) \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix} + \sum_{j\in J} \bar{\mu}_j \begin{pmatrix} D^\top g_j(\bar{x}) \\ 0 \\ \vdots \\ -\bar{z}_j \\ \vdots \\ 0 \end{pmatrix},$$

where $\bar{\mu}_j = 0,\ j \in J \setminus J_0(\bar{x})$. *On the other hand, if* $(\bar{x}, \bar{z})$ *a critical point for* $f|_{\widetilde{M}}$ *, then* $\bar{x}$ *is a critical point for* $f|_M$ *, and* $J_0(\bar{x}) = \{j \in J \,|\, \bar{z}_j = 0\}$.

As a consequence we obtain as our first main result that the number of critical points of $\widetilde{P}$ grows exponentially in the number of inactive constraints, compared to the number of critical points of P.

THEOREM 3.7. *To each critical point* $\bar{x}$ *of* $f|_M$ *correspond exactly* $2^{|J\setminus J_0(\bar{x})|}$ *critical points (and hence KKT-points) of* $f|_{\widetilde{M}}$ *.*

In the next result we show how the linear and quadratic indices of non-degenerate critical points of P and $\widetilde{P}$ are related.

THEOREM 3.8. *Let* $\bar{x} \in M$ *be a non-degenerate critical point for* $f|_M$ *with indices (LI,LCI,QI,QCI). Then* $(\bar{x}, \bar{z})$ *with* $\bar{z}_j = \pm\sqrt{2g_j(\bar{x})}$, $j \in J$, *is a non-degenerate critical point for* $f|_{\widetilde{M}}$ *with indices (0,0,LI+QI,LCI+QCI).*

Proof. Let L denote the Lagrange function corresponding to the non-degenerate critical point $\bar{x}$ of $f|_M$ (cf. (1)), and let $\widetilde{L}$ denote the Lagrange function corresponding to the critical point $(\bar{x}, \bar{z})$ of $f|_{\widetilde{M}}$, i.e.

$$\widetilde{L}(x, z) = f(x) - \sum_{i\in I} \bar{\lambda}_i\, h_i(x) - \sum_{j\in J} \bar{\mu}_j\, (g_j(x) - \tfrac{1}{2} z_j^2)$$

with $\bar{\mu}_j = 0$ for $j \in J \setminus J_0(\bar{x})$. The Hessian of $\widetilde{L}$ possesses the block structure

$$D^2 L(\bar{x}, \bar{z}) = \begin{pmatrix} D^2 L(\bar{x}) & 0 \\ 0 & \operatorname{diag}(\bar{\mu}_j) \end{pmatrix},$$

so that Lemma 3.4 implies that the restricted Hessian $D^2\widetilde{L}(\bar{x}, \bar{z})|_{T_{(\bar{x},\bar{z})}\widetilde{M}}$ has LI+QI negative and LCI+QCI positive eigenvalues. □

REMARK 3.9. As a consequence of Theorem 3.8, the KKT-points of $f|_{\widetilde{M}}$ are not only generated by the KKT-points of $f|_M$ and $(-f)|_M$. In fact, *all* critical points of $f|_M$ give rise to KKT-points for $f|_{\widetilde{M}}$.

By means of an illustrative example let us now investigate in dimension 2 how the Euler characteristic χ (cf. [3]) increases when M is lifted to $\widetilde{M}$.

THEOREM 3.10. *Let $M \subset \mathbb{R}^2$ be a two-dimensional polytope with p vertices ($p \geqslant 3$), defined by p affine linear inequality constraints*

$$0 \leqslant g_j(x) = a_j^\top x - b_j , \; j = 1, ..., p .$$

Then the corresponding Euler characteristics satisfy

$$\chi(M) = 1 \quad \textit{and} \quad \chi(\widetilde{M}) = (4 - p)\, 2^{p-2} .$$

REMARK 3.11. If $p = 3$ then M is a triangle, and $\widetilde{M}$ is a two-dimensional sphere. If $p = 4$, then $\widetilde{M}$ is a torus. In general, the genus (cf. [1]) of the two-dimensional manifold $\widetilde{M}$ equals $(p - 4)\, 2^{p-3} + 1$.

Proof of Theorem 3.10. Choose $c \in \mathbb{R}^2$ such that the function $f(x) = c^\top x$ has only non-degenerate critical points for M. Note that $f|_M$ has exactly one minimum and one maximum, and that at the remaining $p - 2$ critical points it is LI=LCI=1. From Theorem 3.8 it follows that $f|_{\widetilde{M}}$ has 2^{p-2} minima, 2^{p-2} maxima, and $(p - 2)\, 2^{p-2}$ saddle points with QI=QCI=1. Recall that M and, hence, $\widetilde{M}$ is compact (Lemma 3.3). Then, from Morse theory (cf. [3]) we know that $\chi = \sum_{i=0}^{m} (-1)^i c_i$, where c_i denotes the number of KKT-points with QI=i, and where m is the dimension of the manifold (perhaps with boundary). For $f|_M$ we have

$$c_0 = 1 , \quad c_1 = c_2 = 0 ,$$

and for $f|_{\widetilde{M}}$ is is

$$c_0 = c_2 = 2^{p-2} , \quad c_1 = (p - 2)\, 2^{p-2} .$$

This proves the formulas for the corresponding Euler characteristics. □

4. An interior point approach

Throughout this section let Assumption 3.5 (LICQ) hold as well as the following compactness and non-degeneracy assumption.

ASSUMPTION 4.1. *The set M is compact and all KKT-points for $f|_M$ and $(-f)|_M$ are non-degenerate.*

In the following we denote the relative interior of M by

$$M^{>} = \{ x \in \mathbb{R}^n |\; h_i(x) = 0, \; i \in I, \; g_j(x) > 0, \; j \in J \} ,$$

and we put

$$G(x) = \sum_{j \in J} \ln g_j(x)$$

as well as

$$M^\varepsilon \;=\; \{\; x \in M^> \,|\; G(x) \;\geqslant\; \ln\varepsilon \;\}\;.$$

The approach of the present section consists in equalizing the inequality constraint of the approximate set M^ε, i.e. we put

$$\widetilde{M}^\varepsilon \;=\; \{\; (x,z) \in M^> \times \mathbb{R} |\; G(x) \;=\; \ln\varepsilon + \tfrac{1}{2}z^2 \;\}\;.$$

LEMMA 4.2. (cf. [4]).

(i)The set M^ε converges (in the Hausdorff metric) to the set M as $\varepsilon > 0$ tends to zero.

For sufficiently small $\varepsilon > 0$ we have:

(ii)The vectors $Dh_i(x),\ i \in I,\ DG(x)$, are linearly independent for all $x \in M$ satisfying $G(x) = \ln\varepsilon$.
(iii)The set M^ε is homeomorphic with the set M.

Note that for sufficiently small $\varepsilon > 0$ the set $\widetilde{M}^\varepsilon$ is a compact C^2-manifold without boundary and $\dim \widetilde{M}^\varepsilon = \dim M$.

By Assumptions 3.5 and 4.1 the set

$$S \;=\; \{\; x \in M|\; x \text{ is KKT-point for } f|_M \text{ or for } (-f)|_M \;\}$$

is finite.

LEMMA 4.3. ([2, 4]). *For each $x \in S$ let U_x be an (arbitrarily small) neighborhood of x. Then for sufficiently small $\varepsilon > 0$ we have:*

(i)Each critical point of $f|_{M^\varepsilon}$ is non-degenerate, and it belongs to some U_x with $x \in S$.
(ii)There is a one-to-one correspondence between the set S and the set of critical points of $f|_{M^\varepsilon}$. In particular, if $\bar{x} \in M$ is a KKT-point for $f|_M$ (resp. $(-f)|_M$), then the corresponding critical point $\bar{y}$ for $f|_{M^\varepsilon}$ (resp. $(-f)|_{M^\varepsilon}$) is a KKT-point of $f|_{M^\varepsilon}$ (resp. $(-f)|_{M^\varepsilon}$), and the quadratic indices QI at $\bar{x}$ and $\bar{y}$ coincide.

A combination of Theorem 3.7 and Lemma 4.3 immediately yields the following result.

THEOREM 4.4. *Let $\varepsilon > 0$ be sufficiently small. Then the number of critical points for $f|_{\widetilde{M}^\varepsilon}$ is bounded by $2|S|$.*

Hence, the number of critical points does not grow exponentially like in Theorem 3.7 (compare also Remark 3.9). Accordingly the topological complexity of $\widetilde{M}^\varepsilon$ does not change drastically, compared to M. Let us illustrate the latter fact by an example.

THEOREM 4.5. *Let $M \subset \mathbb{R}^n$ be a polytope with non-empty interior, defined by affine linear inequality constraints*

$$0 \leqslant g_j(x) = a_j^\top x - b_j\ , \ j \in J\ .$$

Suppose that LICQ is satisfied at all points of M (Assumption 3.5). Then, for $\varepsilon > 0$ sufficiently small, $\widetilde{M}^\varepsilon$ is homeomorphic to the $n-$dimensional sphere S^n.

REMARK 4.6. Note that M is homeomorphic to the $n-$dimensional Euclidean ball D^n. Hence, from a topological point of view, the transition from M to $\widetilde{M}^\varepsilon$ is just "doubling" the ball D^n into the sphere S^n. Compare this result with Theorem 3.10. Under the assumptions of the latter theorem $\widetilde{M}^\varepsilon$ is just homeomorphic with the two-dimensional sphere S^2.

Proof of Theorem 4.5. Choose $c \in \mathbb{R}^n$ such that the function $f(x) = c^\top x$ has exactly one minimum and one maximum point on M. Then, for sufficiently small $\varepsilon > 0$, $f|_{\widetilde{M}^\varepsilon}$ has exactly two critical points: a non-degenerate minimum and a non-degenerate maximum point. From Morse theory (cf. [3]) we see that the set $\widetilde{M}^\varepsilon$ is built up by two $n-$dimensional cells, homeomorphically glued together along their boundaries. This obviously results into a homeomorphic copy of the $n-$dimensional sphere S^n. □

REMARK 4.7. The assumption of LICQ in Theorem 4.5 can be deleted. In fact, the only technical point consists in checking that the logarithmic smoothing results into a problem with exactly two non-degenerate critical points for $f|_{\widetilde{M}^\varepsilon}$.

5. Final remarks

We point out that in the exact approach (cf. Section 3) as well as in the interior point approach (cf. Section 4) one can consider gradient flows on $\widetilde{M} \subset \mathbb{R}^n \times \mathbb{R}^{|J|}$ and $\widetilde{M}^\varepsilon \subset \mathbb{R}^n \times \mathbb{R}$, respectively, corresponding to a given Riemannian metric on these higher-dimensional spaces. An orthogonal projection of these flows to $\mathbb{R}^n \supset M$ provides ascent and descent differential equations on M.

In [5] we show that this observation yields an automatic adaptation of a given metric on $\mathbb{R}^n$, such that the so-called min-max digraph becomes connected: it is then possible to reach *all* local maxima and minima by integration of the ascent and descent differential equations. The effects of an appropriate *numerical* integration for these differential equations will be subject of future research.

References

1. Griffiths P. and Harris J. (1978), *Principles of Algebraic Geometry,* Wiley, New York.
2. Guerra Vasquez F., Günzel H. and Jongen H.Th. (2001) *On logarithmic smoothing of the maximum function,* Annals of Operations Research, 101, 209–220.
3. Jongen H.Th., Jonker P. and Twilt F. (2000), *Nonlinear Optimization in Finite Dimensions,* Kluwer Academic Publishers, Dordrecht.
4. Jongen H.Th. and Ruiz Jhones A. (1999), *Nonlinear optimization: on the min-max digraph and global smoothing,* in: Ioffe A., Reich S. and Shafrir I. (eds.): Calculus of Variations and Differential Equations, Chapman & Hall/CRC Research Notes in Mathematics Series, 410, CRC Press, Boca Raton, FL, 119–135.
5. Jongen H. Th. and Stein O. (2003), *Constrained global optimization: adaptive gradient flows,* Preprint No. 101, Department of Mathematics - C, Aachen University.
6. Rapcsák T. (1997), *Smooth Nonlinear Optimization in* $\mathbb{R}^n$, Kluwer Academic Publishers, Dordrecht.
7. Tanabe K. (1980), *A geometric method in nonlinear programming,* Journal of Optimization Theory and Applications, 30, 181–210.