Noname manuscript No.
(will be inserted by the editor)

Necessary Global Optimality Conditions for Nonlinear Programming Problems with Polynomial Constraints

V. Jeyakumar · G. Y. Li

Revised Version as a Short Communication: January 16, 2009

Abstract In this paper, we develop necessary conditions for global optimality that apply to non-linear programming problems with polynomial constraints which cover a broad range of optimization problems that arise in applications of continuous as well as discrete optimization. In particular, we show that our optimality conditions readily apply to problems where the objective function is the difference of polynomial and convex functions over polynomial constraints, and to classes of fractional programming problems. Our necessary conditions become also sufficient for global optimality for polynomial programming problems. Our approach makes use of polynomial over-estimators and a powerful theorem of the alternative for a system of polynomials from real algebraic geometry. We discuss numerical examples to illustrate the significance of our optimality conditions.

Keywords Necessary global optimality conditions · polynomial constraints · polynomial over-estimators · algebraic geometry, sums of squares polynomials

1 Introduction

Consider the nonlinear programming problem with polynomial constraints

$$(P)\ \min f(x) \text{ s.t. } g_i(x) \geq 0 \ \ i = 1, \ldots, k, \ h_j(x) = 0 \ \ j = 1, \ldots, l,$$

where f is a twice continuously differentiable function on $\mathbb{R}^n$, g_i, h_j are real polynomials on $\mathbb{R}^n$. Model problems of this form (P) cover a broad range of optimization problems that arise in important applications of continuous as well as discrete optimization. In particular, the model problem (P) encompasses polynomial optimization

Research was partially supported by a grant from the Australian Research Council.

V. Jeyakumar
Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia.
E-mail: v.jeyakumar@unsw.edu.au

G. Y. Li
Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia.
E-mail: g.li@unsw.edu.au

problems, non-convex quadratic programming problems as well as discrete optimization problems such as zero-one nonlinear programming problems where the discrete constraints can be expressed as polynomial equality constraints. In recent years, a great deal of attention has been focused on deriving global optimality conditions for polynomial optimization problems (P) [4,10,15,16], including quadratic programming problems [5,6,8] and bivalent quadratic programming problems [14,9].

A theorem of the alternative is a key tool that is often used in combination with an approximation scheme [1,7] for developing local optimality conditions in continuous optimization. When it comes to deriving global optimality conditions for non-convex optimization problems including certain classes of discrete optimization problems, over or under-estimators of the objective function [8] play an essential role.

Employing a powerful theorem of the alternative from real algebraic geometry [3,11,16,17] together with a polynomial over-estimator, we develop global optimality conditions that don't just provide necessary optimality conditions for nonlinear programming problems (P) with polynomial constraints, but give necessary and sufficient optimality conditions for polynomial programming problems (P). In particular, we show that our necessary conditions readily apply to classes of problems (P) where the objective function is the difference of polynomial and convex functions or a kind of non-convex fractional function.

The outline of this paper is as follows. In section 2, we present necessary global optimality conditions for nonlinear programming problems (P) with polynomial constraints and deduce necessary and sufficient optimality conditions for polynomial programming problems (P). In section 3, we provide qualification-free global optimality conditions for the difference of polynomial and convex minimization problems and fractional programming problems [2]. We provide a numerical example to illustrate that a global minimizer satisfies our necessary conditions whereas a local minimizer that is not global may fail to satisfy them.

2 Necessary Global Optimality Conditions

We begin by fixing notation and preliminaries that will be used later in the paper. As usual, we denote the real polynomial ring on $\mathbb{R}^n$ by $\mathbb{R}[x]$ where $x = (x_1, \ldots, x_n)$. The set of all natural numbers is denoted by $\mathbb{N}$. The notation $A \succeq 0$ means that the matrix A is positive semi-definite. We say that f is a sum of squares polynomial (sos-polynomial) in $\mathbb{R}[x]$ if there exist $k \in \mathbb{N}$, $f_j \in \mathbb{R}[x]$, $j = 1, \ldots, k$ such that for each $x \in \mathbb{R}^n$, $f(x) = \sum_{j=1}^{k} f_j^2(x)$. Let $s \in \mathbb{N}$ and let $f_1, \ldots, f_s \in \mathbb{R}[x]$. We will use the following notation throughout the paper: $S\langle f_1, \ldots, f_s\rangle := \{\prod_{j=1}^{s} f_j^{e_j} : e_j \in \{0,1\}, j = 1, \ldots, s\}$, $I\langle f_1, \ldots, f_s\rangle = \{\sum_{j=1}^{s} a_j f_j : a_j \in \mathbb{R}[x]\}$, $M\langle f_1, \ldots, f_s\rangle = \{\prod_{j=1}^{s} f_j^{e_j} : e_j \in \mathbb{N} \cup \{0\}\}$ and $C\langle f_1, \ldots, f_s\rangle := \{\sum_{i=1}^{2^s} y_i u_i : y_i \text{ is a sos-polynomial}, u_i \in S\langle f_1, \ldots, f_s\rangle, i = 1, \ldots, 2^s\}$.

The following Theorem of the Alternative (cf. [16, Theorem 2.2]) for polynomial systems plays an important role in deriving our main result, Theorem 1.

Lemma 1 *(Theorem of the Alternative) Let $f \in \mathbb{R}[x]$. Let $k, l \in \mathbb{N}$ and let $g_i, h_j \in \mathbb{R}[x]$, $i = 1, \ldots, k$ and $j = 1, \ldots, l$. Then exactly one of the following holds:*
(1) There exists some x_0 such that $g_i(x_0) \geq 0$, $h_j(x_0) = 0$ and $f(x_0) > 0$.
(2) There exists $z_1 \in C\langle g_1, \ldots, g_k, f\rangle$, $z_2 \in I\langle h_1, \ldots, h_l\rangle$ and $z_3 \in M\langle f\rangle$ such that, for each $x \in \mathbb{R}^n$, $z_1(x) + z_2(x) + z_3(x) = 0$.

Recall that a function $h : \mathbb{R}^n \to \mathbb{R}$ is an *over-estimator* of the function $f : \mathbb{R}^n \to \mathbb{R}$ at $\bar{x}$ over a set $\Delta \subset \mathbb{R}^n$ if, for each $x \in \Delta$, $f(x) \leq h(x)$, and $f(\bar{x}) = h(\bar{x})$. The function h is said to be a *polynomial over-estimator* if $h \in \mathbb{R}[x]$ and is an over-estimator of f at $\bar{x}$ over a set Δ. The feasible set of (P) is given by $F := \{x \ : \ g_i(x) \geq 0,\ h_j(x) = 0,\ i = 1, \ldots, k\ j = 1, \ldots, l\}$. For $x_0 \in \mathbb{R}^n$, we define the functions g, $\tilde{g}_{x_0} : \mathbb{R}^n \to \mathbb{R}$ by $g(x) := p(x) - (\nabla p(x_0) - \nabla f(x_0))^T x$ and $\tilde{g}_{x_0}(x) = g(x_0) - g(x)$.

By employing a polynomial outer-estimator and Lemma 1 we derive a general global optimality conditions for a broad class of non-convex problems (P).

Theorem 1 *For the problem (P), suppose that there exist a convex set $C \supseteq F$ and $p \in \mathbb{R}[x]$ such that, for each $x \in C$, $\nabla^2 f(x) - \nabla^2 p(x) \preceq 0$. If x_0 is a global minimizer of (P), then there exist $\gamma \in \mathbb{N}$, $a_j \in \mathbb{R}[x]$, $j = 1, \ldots, l$, sos-polynomials y_i in $\mathbb{R}[x]$ and $u_i \in S\langle g_1, \ldots, g_k, \tilde{g}_{x_0}\rangle$, $i = 1, \ldots, 2^{k+1}$ such that* $\min\{y_i(x_0), u_i(x_0)\} = 0$ *for $i \in \{1, \ldots, 2^{k+1}\}$ and, for each $x \in \mathbb{R}^n$,*

$$\sum_{i=1}^{2^{k+1}} y_i(x)u_i(x) + \sum_{j=1}^{l} a_j(x)h_j(x) + [\tilde{g}_{x_0}(x)]^\gamma = 0. \tag{1}$$

Proof Let g be defined as above and $\phi(x) := f(x) - g(x)$. Then, for each $x \in C$, $\nabla^2 \phi(x) = \nabla^2 f(x) - \nabla^2 p(x) \preceq 0$. Consequently, ϕ is concave on C. By construction, its gradient vanishes at x_0. So, the function h, defined by $h(x) = g(x) - g(x_0) + f(x_0)$, is a polynomial over-estimator of f at x_0 over C. This gives us that for each $x \in F \subseteq C$, $f(x_0) - f(x) \geq \tilde{g}_{x_0}(x)$. Since x_0 is a global minimizer of (P), it follows that the system of inequalities $g_i(x) \geq 0, i = 1, \ldots, k, h_j(x) = 0, j = 1, \ldots, l, \tilde{g}_{x_0}(x) > 0$ is inconsistent. Now, by Lemma 1 with f replaced by $\tilde{g}_{x_0}$, we obtain that: there exist $\gamma \in \mathbb{N} \cup \{0\}$, $z_1 \in C\langle g_1, \ldots, g_k, \tilde{g}_{x_0}\rangle$ and $z_2 \in I\langle h_1, \ldots, h_l\rangle$ such that, for each $x \in \mathbb{R}^n$, $z_1(x) + z_2(x) + \tilde{g}^\gamma_{x_0}(x) = 0$. Therefore, there exist $\gamma \in \mathbb{N} \cup \{0\}$, $a_j \in \mathbb{R}[x]$, $j = 1, \ldots, l$, sos-polynomials y_i and $u_i \in S\langle g_1, \ldots, g_k, \tilde{g}_{x_0}\rangle$, $i = 1, \ldots, 2^{k+1}$ such that, for each $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{2^{k+1}} y_i(x)u_i(x) + \sum_{j=1}^{l} a_j(x)h_j(x) + [\tilde{g}_{x_0}(x)]^\gamma = 0. \tag{2}$$

Substituting $x = x_0$ in (2) and noting that $\tilde{g}_{x_0}(x_0) = 0$, $y_i's$ are sums of squares (hence nonnegative) and $x_0 \in F$ (and hence $u_i(x_0) \geq 0$, $h_j(x_0) = 0$), we obtain that $\min\{y_i(x_0), u_i(x_0)\} = 0$ for each $i \in \{1, \ldots, 2^{k+1}\}$. Note that $\gamma \neq 0$. Otherwise, $0 \leq \sum_{i=1}^{2^{k+1}} y_i(x_0)u_i(x_0) + \sum_{j=1}^{l} a_j(x_0)h_j(x_0) = -1$, which is a contradiction.

The following example shows that, in general, the requirement of $\gamma > 0$ in Theorem 1 cannot be strengthened to $\gamma = 1$.

Example 1 Consider the following one dimensional minimization problem

$$\min\ x - 2e^{-x} \text{ s.t. } x^3 \geq 0.$$

Clearly, $x_0 = 0$ is a global minimizer of the problem. Take $p(x) = x$. Then our necessary condition holds with $\gamma = 3$ since $27 \cdot x^3 + (-3x)^3 = 0$. However, this condition does not hold with $\gamma = 1$, since for any sos-polynomials y_i and $u_i \in S\langle x^3, -3x\rangle$, $i = 1, 2, 3, 4$, $\sum_{i=1}^{4} y_i(x)u_i(x) + (-3x)$ is not the zero polynomial.

Polynomial Programming Problems

We see that our necessary condition becomes also sufficient for problem (P) whenever f is a polynomial (cf. [16]). Indeed, in this case, if we choose the polynomial over-estimator p as the objective function f then (1) becomes $\sum_{i=1}^{2^{k+1}} y_i(x)u_i(x) + \sum_{j=1}^{l} a_j(x)h_j(x) + [f(x_0) - f(x)]^\gamma = 0$. So, for each feasible point $x \in F$, $[f(x_0) - f(x)]^\gamma = -\sum_{i=1}^{2^{k+1}} y_i(x)u_i(x) - \sum_{j=1}^{l} a_j(x)h_j(x) \leq 0$. This shows that $f(x_0) \leq f(x)$ for each feasible point $x \in F$. Thus, (1) is sufficient to global optimality of (P). As noted in [16], checking our necessary condition (1) can be reduced to solving a sequence of (convex) semidefinite programs.

Zero-One Programming Problems

It is worth noting that Theorem 1 applies to nonlinear programming problems with 0/1 constraints, $x_j \in \{0, 1\}$, since the 0/1 constraints can be equivalently rewritten as the polynomial equality constraints $x_j^2 - x_j = 0$.

Smooth Programming Problems with Bounded Feasible Sets

Let us assume that the feasible set F of (P) is bounded. Let the Hessian of f be defined by $\nabla^2 f(x) = \Big(f_{ij}(x)\Big)$ and let C be a convex and compact set containing F. Define a_i, $i = 1, \ldots, n$ and $p : \mathbb{R}^n \to \mathbb{R}$ by $a_i = \max\{f_{ii}(x) + \sum_{j \neq i,\, 1 \leq j \leq n} |f_{ij}(x)| : x \in C\}$ [1] and $p(x) = \frac{1}{2}\sum_{i=1}^{n} a_i x_i^2$. As illustrated in [6], we can easily show that $\nabla^2 f(x) - \nabla^2 p(x) \preceq 0$ for each $x \in C$. So, Theorem 1 applies to problem (P) with bounded feasible set F.

3 Qualification-Free Global Optimality Conditions

In this Section, we present qualification-free necessary global optimality conditions for some important classes of problems. We begin by examining the minimization of the difference of polynomial and convex functions over polynomial constraints.

Difference of Polynomial and Convex Minimization

Consider the following difference of polynomial and convex minimization problem

$$(P_1)\ \min\ p(x) - k(x)\ \text{s.t.}\ g_i(x) \geq 0,\ i = 1, \ldots, k,\ \ h_j(x) = 0,\ j = 1, \ldots, l,$$

where p is a real polynomial on $\mathbb{R}^n$ and k is a twice continuously differentiable convex function on $\mathbb{R}^n$, g_i, h_j are real polynomials on $\mathbb{R}^n$. The model problem (P_1) serves as an extension of the polynomial optimization and concave minimization. We now present a necessary global optimality condition for problem (P_1). We define $\tilde{g}^*_{x_0} : \mathbb{R}^n \to \mathbb{R}$ by $\tilde{g}^*_{x_0}(x) := [p(x_0) - p(x)] - \nabla k(x_0)^T (x_0 - x)$.

Corollary 1 *Let x_0 be a global minimizer of (P_1). Then there exist $\gamma \in \mathbb{N}$, $a_j \in \mathbb{R}[x]$ $j = 1, \ldots, l$, sos-polynomials y_i in $\mathbb{R}[x]$ and $u_i \in S\langle g_1, \ldots, g_k, \tilde{g}^*_{x_0}\rangle$, $i = 1, \ldots, 2^{k+1}$, such that* $\min\{y_i(x_0), u_i(x_0)\} = 0$ *for each* $i \in \{1, \ldots, 2^{k+1}\}$ *and, for each* $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{2^{k+1}} y_i(x)u_i(x) + \sum_{j=1}^{l} a_j(x)h_j(x) + [\tilde{g}^*_{x_0}(x)]^\gamma = 0. \quad (3)$$

[1] Calculating the constant a_i's is, in general, difficult. However, if the objective function f is a separable function, i.e $f(x) = \sum_{i=1}^{n} f_i(x_i)$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then we can choose C as a bounded box containing F, and each a_i can be efficiently calculated by solving a one-dimensional maximization problem.

Proof Let $f(x) = p(x) - k(x)$ for each $x \in \mathbb{R}^n$. Then, $\nabla^2 f(x) - \nabla^2 p(x) = -\nabla^2 k(x) \preceq 0$. Moreover, for each $x \in \mathbb{R}^n$,

$$\tilde{g}_{x_0}(x) := [p(x_0) - p(x)] - (\nabla p(x_0) - \nabla f(x_0))^T (x_0 - x) = \tilde{g}^*_{x_0}(x).$$

Hence, the conclusion follows immediately from Theorem 1.

We now present an example to illustrate that a global minimizer satisfies our necessary condition whereas a local minimizer that is not global fails to satisfy the necessary condition.

Example 2 Consider the minimization problem

$$(E2)\ \min x_1^2 - 3x_2^2 - e^{x_2} \quad \text{s.t.} \quad -x_1^2 - x_2^2 + 9 \geq 0,\ 3 - x_2 \geq 0,\ x_1 - 1 \geq 0,\ x_1 x_2 = 0.$$

The feasible set $F = \Big(\{0\} \times [-3, 3]\Big) \cup \Big([1, 3] \times \{0\}\Big)$. It is easy to check that $x_0 = (0, 3)$ is the global minimizer of problem (E2). Moreover, it can be verified that $z_0 = (1, 0)$ is a local minimizer. Let $f(x) = x_1^2 - 3x_2^2 - e^{x_2}$, $g_1(x) = -x_1^2 - x_2^2 + 9$, $g_2(x) = 3 - x_2$, $g_3(x) = x_1 - 1$ and $h(x) = x_1 x_2$. Then $f(x) = p(x) - k(x)$ where $p(x) = x_1^2 - 3x_2^2$ and $k(x) = e^{x_2}$. Clearly, p is a real polynomial and k is a twice continuously differentiable convex function. Moreover, for each $x = (x_1, x_2)$ one has $\tilde{g}^*_{x_0}(x) = [-27 - x_1^2 + 3x_2^2] + e^3(x_2 - 3) = -x_1^2 + 3x_2^2 + e^3 x_2 - 27 - 3e^3$. Note that the following polynomial equality $4x_1^2 + 3g_1(x) + e^3 g_2(x) + \tilde{g}^*_{x_0}(x) = 0$ holds for each $x \in \mathbb{R}^n$. It now follows that our necessary condition holds at the point x_0. On the other hand, it can be verified that $\tilde{g}^*_{z_0}(x) = x_1^2 + 3x_2^2 + x_2$. To see our necessary condition fails for z_0, we suppose on the contrary that there exist $\gamma \in \mathbb{N}$, real polynomial a, sos-polynomial y_i, and $u_i \in S\langle g_1, g_2, g_3, \tilde{g}^*_{z_0}\rangle$ $(i = 1, \ldots, 16)$ such that $\sum_{i=1}^{16} y_i(x) u_i(x) + a(x) h(x) + [\tilde{g}^*_{z_0}(x)]^\gamma \equiv 0$. Substituting $x = x_0$, as $h(x_0) = 0$, we get that $\sum_{i=1}^{16} y_i(x_0) u_i(x_0) + [\tilde{g}^*_{z_0}(x_0)]^\gamma = 0$. Note that $\tilde{g}^*_{z_0}(x_0) > 0$, $y_i(x_0) \geq 0$ and $u_i(x_0) \geq 0$. This is a contradiction.

Fractional Programming Problems

Consider the following fractional minimization problem

$$(FP)\ \min \tfrac{v(x)}{p(x)} \text{ s.t. } g_i(x) \geq 0,\ i = 1, \ldots, k,\ h_j(x) = 0,\ j = 1, \ldots, l,$$

where p is a positive real polynomial function on $\mathbb{R}^n$, v is a twice continuously differentiable concave function and g_i, h_j are real polynomials on $\mathbb{R}^n$. Note that x_0 is a global minimizer of (FP) if and only if x_0 is a global minimizer of the following minimization problem

$$(P')\ \min(-\tfrac{v(x_0)}{p(x_0)})p(x) - (-v(x)) \text{ s.t. } g_i(x) \geq 0,\ i = 1, \ldots, k,\ h_j(x) = 0,\ j = 1, \ldots, l.$$

Now, by applying Corollary 1 we can easily derive necessary global optimality conditions for (FP).

4 Conclusion and Future Research

Employing a powerful theorem of the alternative from real algebraic geometry together with a polynomial over-estimator, we developed necessary conditions for global optimality that apply to broad classes of non-convex minimization problems (P). It would be of practical interest to examine how (and for what class of problems) our necessary conditions can efficiently be checked. Example 2 illustrated that our optimality conditions may be used to identify global minimizers for certain problems. However, it would be helpful to refine our optimality conditions so that they can be used, in general, to rule out local minimizers that are not global. On the other hand, a method of checking our necessary condition (1) involves solving a hierarchy of semidefinite programming problems. However, due to numerical inaccuaracies inherent to a semidefinite programming solver, such a scheme would be practically difficult in obtaining the certificate (1). The issues associated with the practical and numerical aspects of our optimality conditions will be investigated in a further study.

Acknowledgements Authors are grateful to the referees for their helpful comments and valuable suggestions which have contributed to the final preparation of the paper.

References

1. M. Bazaraa, H. Sherali, and C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*, John Wiley. Sons, (1993).
2. H. P. Benson, Fractional programming with convex quadratic forms and functions, *European Journal of Operational Research*, 173 (2006), 351-369.
3. J. Bochnak, M. Coste and M. F. Roy, *Real Algebraic Geometry*, Springer-Verlag, (1998).
4. D. Henrion and J. B. Lasserre, Detecting global optimality and extracting solutions in GloptiPoly, Positive polynomials in control, 293-310, Lecture Notes in Control and Inform. Sci., 312, Springer, Berlin, (2005).
5. J. B. Hiriart-Urruty, Conditions for Global Optimality 2, *J. Global Optim.*, 13 (1998), 349-367.
6. V. Jeyakumar and S. Srisatkunarajah, Lagrange multiplier necessary conditions for global optimality for non-convex minimizations over a quadratic constraint via S-lemma, *Optim. Lett.*, 3 (2009), No. 1, 23-33.
7. V. Jeyakumar and D. T. Luc, *Nonsmooth Vector Functions and Continuous Optimization* , Vol. 10, Series on Optimization and Its Applications Series, Springer, New York, (2008).
8. V. Jeyakumar, A. M. Rubinov and Z. Y. Wu, Sufficient global optimality conditions for non-convex quadratic minimization problems with box constraints, *J. Global Optim.*, 36 (2006), 471-481.
9. V. Jeyakumar, A. M. Rubinov and Z. Y. Wu, Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions, *Math. Program, Series A*, 110 (2007), 521-541.
10. J. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM. J. Optimization*, 11 (2001), 796-817.
11. M. Marshall, *Positive Polynomials and Sums of Squares*, AMS. Math. Surveys and Monographs, 146 (2008).
12. J. W. Nie, J. Demmel and B. Sturmfels, Minimizing polynomials via sum of squares over the gradient ideal, Math. *Program, Series A*, 106 (2006), 587-606.
13. J. W. Nie, Demmel and V. Powers, Representations of positive polynomials on non-compact semialgebraic sets via KKT ideals, *J. Pure Appl. Algebra* 209 (2007), no. 1, 189-200..
14. P. Pardalos and H. Romeijn, *Handbook in Global Optimization*, Kluwer academic publishers, (2002).
15. M. Putinar, Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* 42 (1993), no. 3, 969-984.
16. H. Schichl and A. Neumaier, Transposition theorems and qualification-free optimality conditions, *SIAM. J. Optimization* 17 (2006), 1035-1055.
17. G. Stengle, A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Math. Ann.* 207 (1974), 87-97.