Bull. Math. Soc. Sci. Math. Roumanie
Tome 49(97) No. 3, 2006, 239–245

Local Convexification of the Lagrangian Function for Nonlinear Nonconvex Optimization

by
CRISTINCA FULGA

Abstract

In this paper, we are concerned with a differentiable nonlinear programming problem with inequality constraints. By using an equivalent transformation of the constrained optimization problem we establish the local convexity of the Lagrangian function of the transformed equivalent problem. Zero duality gap is thus guaranteed when the primal-dual method is applied to the constructed equivalent form.

Key Words: Nonconvex Optimization, Lagrangian Function, Local Convexification.
2000 Mathematics Subject Classification: Primary 90C30, Secondary 90C26.

1 Introduction

Convexity plays a vital role in many aspects of mathematical programming including optimality conditions and duality theorems. To relax convexity assumptions imposed on the functions in theorems on optimality and duality, various generalized convexity concepts have been proposed. Important contributors in this field are Preda (see for example[8]), Hanson ([2], [9]), Jeyakumar and Mond ([3]).

The celebrated primal-dual method (see for example Lasdon [4], Luenberger [7]) has been one of the most efficient solution algorithms in solving constrained optimization problems under certain conditions. The primal-dual method resorts to the convergence via sequential minimization of the Lagrangian function. The succes of the primal-dual method depends on the local convexity of the Lagrangian function at the optimal solution of the problem. Several convexification schemes have been proposed in the literature to extend the primal- dual method for certain nonconvex problems, see for example Bertsekas [1], Li [5] and [6], Xu [10].

In this paper we consider the following general constrained optimization programming problem in nonlinear programming:

$$\begin{cases} \min f_0(x) \\ \text{s.t. } f_j(x) \le b_j,\ j = \overline{1,m} \\ x \in X, \end{cases} \tag{1}$$

where $f_j : R^n \to R$, $j = \overline{0,m}$ are twice continuously differentiable functions and $X \subset R^n$ is a nonempty closed set. The primal-dual method has been one of the most efficient instrument in solving problem 1 under certain conditions. The primal-dual method resorts to the convergence via sequential minimization of the Lagrangian function of 1 defined by

$$L(x,\lambda) = f_0(x) + \sum_{j=1}^{m} \lambda_j [f_j(x) - b_j],\ x \in X,\ \lambda = (\lambda_1, ..., \lambda_m) \in R_+^m. \tag{2}$$

The succes of the primal-dual method has been limited to the case when $L(x,\lambda)$ is locally convex at the optimal solution of problem 1. As revealed in several studies, convexity or nonconvexity is not an inherent property in optimization. A set could be nonconvex in one representation space, while the same set could become convex when changing the coordinates of the representation space, see for example [6].

2 Local Convexity of the Lagrangian Function of the Equivalent Form

Let x^* be a regular point of the constraints in 1 i.e. $\nabla f_j(x^*)$, $j \in I(x^*)$ are linearly independent, where $I(x^*) = \left\{ j \left| f_j(x^*) = b_j,\ j = \overline{1,m} \right. \right\}$. We assume that x^* satisfies the second-order sufficiency condition; then there exists a Lagrange multiplier $\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in R_+^m$. such that

$$\nabla L(x^*, \lambda^*) = 0 \tag{3}$$

$$\lambda_j^* (f_j(x^*) - b_j) = 0, j = \overline{1,m} \tag{4}$$

and the Hessian matrix is positive on the tangent subspace $M(x^*)$,

$$y^T \nabla^2 L(x^*, \lambda^*) y > 0, \forall y \in M(x^*), y \neq 0, \tag{5}$$

where $M(x^*) = \left\{ y \in R^n \left| y^T \nabla f_j(x^*) = 0, j \in J(x^*) \right. \right\}$ and $J(x^*) = \left\{ j \left| \lambda_j^* > 0 \right. \right\}$.

Let φ be a montone real function with $\varphi' > 0$ on R and locally convex in $r^* = f_0(x^*)$, $b_1, b_2, ..., b_m$, where φ' is the derivative function of the real function φ. Let $p \ge 1$ and $q \ge 1$. Further, we assume that φ (and thus, φ^p and φ^q) is positive over R. This assumption does not impose a loss of generality as we can always apply some suitable equivalent transformation if necessary.

We consider the following problem:

$$\begin{cases} \min \varphi^p\left(f_0\left(x\right)\right) \\ \text{s.t. } \varphi^q\left(f_j\left(x\right)\right) \leq \varphi^q\left(b_j\right), \ j = \overline{1,m} \\ x \in X, \end{cases} \tag{6}$$

which is equivalent with problem 1. The Lagrangian function associated with problem 6 is defined by introducing Lagrangian multiplier $\mu = (\mu_1, ..., \mu_m) \in R^m_+$:

$$L_\varphi\left(x, \mu\right) = \varphi^p\left(f_0\left(x\right)\right) + \sum_{j=1}^{m} \mu_j \left[\varphi^q\left(f_j\left(x\right)\right) - c_j\right], \ x \in X \tag{7}$$

where $c_j = \varphi^q\left(b_j\right), \ j = \overline{1,m}$. The vector of optimal multipliers $\mu^* = (\mu_1^*, ..., \mu_m^*)$ associated with x^* in the Lagrangian 7 are given by

$$\mu_j^* = \begin{cases} \lambda_j^* \frac{p}{q} \frac{\varphi^{p-1}(r^*)}{\varphi^{q-1}(b_j)} \frac{\varphi'(r^*)}{\varphi'(b_j)}, \ j \in J\left(x^*\right) \\ 0, \ \text{otherwise.} \end{cases} \tag{8}$$

The Hessian matrix has the form

$$\nabla^2 L_\varphi\left(x^*, \mu^*\right) = p\varphi^{p-1}\left(r^*\right)\varphi'\left(r^*\right)\Big[\nabla^2 L\left(x^*, \lambda^*\right) +$$

$$+\left(u_0 + (p-1)\,v_0\right)\nabla f_0\left(x^*\right)\nabla f_0^T\left(x^*\right) +$$

$$+\sum_{j=1}^{m}\left(u_j + (q-1)\,v_j\right)\nabla f_j\left(x^*\right)\nabla f_j^T\left(x^*\right)\Big],$$

where

$$\begin{cases} u_0 = \frac{\varphi''(r^*)}{\varphi'(r^*)} \\ u_j = \lambda_j^* \frac{\varphi''(b_j)}{\varphi'(b_j)}, \ j \in J\left(x^*\right) \\ u_j = 0, \ j \notin J\left(x^*\right) \end{cases}$$

and

$$\begin{cases} v_0 = \frac{\varphi'(r^*)}{\varphi(r^*)} \\ v_j = \lambda_j^* \frac{\varphi'(b_j)}{\varphi(b_j)}, \ j \in J\left(x^*\right) \\ v_j = 0, \ j \notin J\left(x^*\right). \end{cases}$$

Theorem 2.1. Let x^* be a local optimal solution of problem 1. Assume that $J\left(x^*\right)$ is nonempty, x^* is a regular point of the constraints in 1 and x^* satisfies the second order sufficiency condition. Let φ be a real function with $\varphi' > 0$ and locally convex in $r^* = f_0\left(x^*\right)$, $b_1, b_2, ..., b_m$ and $p \geq 1$, $q \geq 1$. Then, there exist $\widetilde{p} > 0$ and $\widetilde{q} > 0$ such that the hessian matrix $\nabla^2 L_\varphi\left(x^*, \mu^*\right)$is positive definite when $p \geq \widetilde{p}$ and $q \geq \widetilde{q}$.

Proof: Due to the assumptions on φ, we have $u_j > 0$ and $v_j > 0$, $j \in \widetilde{J}\left(x^*\right)$, where $\widetilde{J}\left(x^*\right) = J\left(x^*\right) \cup \{0\}$.

First, we consider $M(x^*) = \{0\}$; that means that for all $y \neq 0$ there is at least one $j \in J(x^*)$ such that $y^T \nabla f_j(x^*) \neq 0$. We make the notations:

$$\tau_1 = \min_{y \in S} \sum_{j \in \widetilde{J}(x^*)} u_j \left[y^T \nabla f_j(x^*) \right]^2, \tag{9}$$

$$\tau_2 = \min_{y \in S} \sum_{j \in J(x^*)} v_j \left[y^T \nabla f_j(x^*) \right]^2 \tag{10}$$

and

$$\eta = \min_{y \in S} y^T \nabla^2 L(x^*, \lambda^*) y \tag{11}$$

where $S = \{ y \in R^n \,|\, \|y\| = 1 \}$. We remark that $\tau_1 > 0$ and $\tau_2 > 0$. For any $y \in S$, from **??**-11 we have

$$y^T \nabla^2 L_\varphi(x^*, \mu^*) y \geq$$

$$\geq p\varphi^{p-1}(r^*) \varphi'(r^*) \left\{ \eta + \tau_1 + (p-1) v_0 \left[y^T \nabla f_0(x^*) \right]^2 + (q-1) \tau_2 \right\}.$$

If $\eta = 0$, then for any $p \geq 1$ and $q \geq 1$ we have

$$y^T \nabla^2 L_\varphi(x^*, \mu^*) y \geq p\varphi^{p-1}(r^*) \varphi'(r^*) \tau_1 > 0.$$

If $\eta < 0$ we consider two cases. First, for those $y \in S$ for which $y^T \nabla f_0(x^*) = 0$ we remark that, for $q \geq 1 - \frac{\eta}{\tau_2}$ we obtain

$$y^T \nabla^2 L_\varphi(x^*, \mu^*) y > 0.$$

Then, for $y \in S$ for which $y^T \nabla f_0(x^*) \neq 0$ we use the notation

$$\tau_3 = \min_{y \in S - \{ y | y^T \nabla f_0(x^*) = 0 \}} v_0 \left[y^T \nabla f_0(x^*) \right]^2.$$

Taking

$$\begin{cases} p \geq \max\left\{1; 1 - \frac{\eta}{\tau_3}\right\} \\ q \geq \max\left\{1; 1 - \frac{\eta}{\tau_2}\right\} \end{cases}$$

the Hessian matrix $\nabla^2 L_\varphi(x^*, \mu^*)$ is positive definite on S and therefore on R^n.

Now, we consider the case $M(x^*) \neq \{0\}$. So, there is $y \neq 0$ with $y^T \nabla f_j(x^*) = 0, \forall\, j \in J(x^*)$, i.e. with $\angle(y, \nabla f_{j_0}(x^*)) = \frac{\pi}{2}, \forall\, j \in J(x^*)$, where $\angle(a, b) \in [0, \pi]$ denotes the angle between $a \in R^n$ and $b \in R^n$. From 5 we have

$$\varepsilon = \min_{y \in M(x^*) \cap S} y^T \nabla^2 L(x^*, \lambda^*) y > 0. \tag{12}$$

For $y \in M(x^*) \cap S$,

$$y^T \nabla^2 L_\varphi(x^*, \mu^*) y \geq$$

$$\geq p\varphi^{p-1}(r^*) \varphi'(r^*) \left\{ \varepsilon + (u_0 + (p-1) v_0) \left[y^T \nabla f_0(x^*) \right]^2 \right\} > 0.$$

For any $y \in S - M(x^*)$ it exists $j_0 \in J(x^*)$ such that $\angle(y, \nabla f_{j_0}(x^*)) \neq \frac{\pi}{2}$.Using 12 and the compactness of $M(x^*) \cap S$ we find that there exists $\theta_0 > 0$ such that for any $y \in S$ with $\angle(y, \nabla f_j(x^*)) \in \left[\frac{\pi}{2} - \theta_0, \frac{\pi}{2} + \theta_0\right]$ we have $y^T \nabla^2 L(x^*, \lambda^*) y \geq \frac{\varepsilon}{2}$. So, $y^T \nabla^2 L_\varphi(x^*, \mu^*) y \geq p\varphi^{p-1}(r^*) \varphi'(r^*) \frac{\varepsilon}{2} > 0$.

For any $y \in S$ for which there exists $j_0 \in J(x^*)$ such that $\angle(y, \nabla f_{j_0}(x^*)) \notin \left[\frac{\pi}{2} - \theta_0, \frac{\pi}{2} + \theta_0\right]$ we make the notations:

$$c = \min\left\{|\cos\theta| \,\middle|\, \left|\theta - \frac{\pi}{2}\right| \geq \theta_0, \theta \in [0, \pi]\right\} \tag{13}$$

$$\delta = \min_{j \in J(x^*)} \|\nabla f_j(x^*)\| \tag{14}$$

$$\widetilde{u} = \min_{j \in J(x^*)} u_j \tag{15}$$

$$\widetilde{v} = \min_{j \in J(x^*)} v_j. \tag{16}$$

$$\psi = \left[y^T \nabla f_0(x^*)\right]^2 \tag{17}$$

We remark that $\delta > 0, \widetilde{u} > 0, \widetilde{v} > 0$ and $\psi \geq 0$. Combining this remark with **??** and 13-17 we get

$$\begin{gathered} y^T \nabla^2 L_\varphi(x^*, \mu^*) y \geq \\ \geq p\varphi^{p-1}(r^*) \varphi'(r^*) \{\eta + (u_0 + (p-1) v_0) \psi + \\ + (\widetilde{u} + (q-1) \widetilde{v}) \delta^2 c^2\}. \end{gathered}$$

We find that $y^T \nabla^2 L_\varphi(x^*, \mu^*) y > 0$ for all $q \geq 1 - \frac{\eta}{\widetilde{v}\delta^2 c^2}$ and $p \geq p(\psi)$where the threshold $p(\psi)$is

$$p(\psi) = \begin{cases} 1, \text{ if } \psi = 0 \\ 1 - \frac{\eta}{v_0 \psi^2}, \text{ if } \psi \neq 0 \end{cases}.$$

□

3 Duality Result

We define the dual function associated with the problem 6:

$$D_\varphi(\mu) = \min_{x \in X} L_\varphi(x, \mu)$$

where here it is understood that the minimum is taken locally with respect to x near x^*.

Theorem 3.1. Suppose that the problem 1 has a local solution at $x^*$with corresponding value r^* and Lagrange multiplier λ^*. Also suppose that $J(x^*)$ is nonempty, x^* is a regular point of the constraints and x^* satisfies the second order sufficiency condition.Let x^* be a local optimal solution of problem 1. Let

φ be a real function with $\varphi' > 0$ on R, locally convex in $r^* = f_0(x^*), b_1, b_2, ...b_m$ and $p \geq 1$, $q \geq 1$. Then, there exist $\widetilde{p} > 0$ and $\widetilde{q} > 0$ such that the dual problem

$$\max_{\mu \geq 0} D_\varphi(\mu)$$

has a local solution μ^* defined by 8 with optimal value $\varphi^p(r^*)$ and $L_\varphi(x^*, \mu^*) = D_\varphi(\mu^*), \forall\ p \geq \widetilde{p}$ and $q \geq \widetilde{q}$.

Proof: Combining the Local Duality Theorem ([7]) and Theorem 2.1. the result is imediat. □

Remark 3.1. *Theorem 3.1. generalizes some results in Xu [10] obtained under slightly different conditions.*

4 Conclusions

In this paper we have shown how to locally convexify the Lagrangian function of a nonconvex optimization problem and thus expand the class of optimization problems to which dual methods can be applied. Specifically, we have proved that, under mild assumptions, the Hessian of the Lagrangian function in some transformed equivalent problem formulations becomes positive definite in a neighborhood of a local optimal point of the original problem. From local duality theory, convexity in the Lagrangian guarantees the succes of the dual search and zero duality gap is thus guaranteed when the primal-dual method is applied to the constructed equivalent form.

References

[1] D. P. Bertsekas, *Convexification Procedures and Decomposition Methods for Nonconvex Optimization Problems.* Journal of Optimization Theory and Applications, **29** (1979), 169-197.

[2] M. Hanson, *On Sufficiency of Kuhn-Tucker Conditions.* Journal of Mathematical Analysis and Applications, **80** (1981), 545-550.

[3] V. Jeyakumar, B. Mond, *On Generalized Convex Mathematical Programming.* Journal of Australian Mathematical Society, Series B, **34** (1992), 43-53.

[4] L. S. Lasdon, *Optimization Theory for Large Systems.* Macmillan Company, London, England, 1970.

[5] D. Li, *Zero Duality Gap for a Class of Nonconvex Optimization Problems.* Journal of Optimization Theory and Applications, **85** (1995), 309-324.

[6] D. Li, *Saddle-Point Generation in Nonlinear Nonconvex Optimization.* Nonlinear Analysis, Theory, Methods and Applications, **30** (1997), 4339-4344.

[7] D. G. Luenberger, *Linear and Nonlinear Programming.* 2nd Edition, Addison-Wesley, Reading, Massachusetts, 1984.

[8] V. Preda, *On Sufficiency and Duality for Generalized Quasiconvex Programs.* Journal of Mathematical Analysis and Applications, **181** (1994), 77-88.

[9] N. G. Rueda, M. A Hanson, C. Singh, *Optimality and Duality with Generalized Convexity.* Journal of Optimization Theory and Applications, **86** (1995), 491-500.

[10] Z. K. Xu, *Local Saddle Points and Convexification for Nonconvex Optimization Problems.* Journal of Optimization Theory and Applications, **94** (1997), 739-746.

Received: 15.03.2006.
Revised: 30.03.2006.

Academy of Economic Studies,
Faculty of Cybernetics, Statistics and Economic Informatics,
Department of Mathematics,
Piaţa Romană, 6
Bucharest, Romania
E-mail: fulga@csie.ase.ro