Research Article

Strong Convergence on the Aggregate Constraint-Shifting Homotopy Method for Solving General Nonconvex Programming

Zhichuan Zhu and Yeong-Cheng Liou

School of Economics, Liaoning University, Shenyang, Liaoning 110136, China
Department of Healthcare Administration and Medical Informatics, and Research Center of Nonlinear Analysis and Optimization, Kaohsiung Medical University, Kaohsiung 807, Taiwan
Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung 807, Taiwan

Correspondence should be addressed to Yeong-Cheng Liou; simplex_liou@hotmail.com

Received 16 September 2020; Revised 20 November 2020; Accepted 22 November 2020; Published 15 December 2020

Academic Editor: Ching-Feng Wen

Copyright © 2020 Zhichuan Zhu and Yeong-Cheng Liou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the paper, the aggregate constraint-shifting homotopy method for solving general nonconvex nonlinear programming is considered. The aggregation is only about inequality constraint functions. Without any cone condition for the constraint functions, the existence and convergence of the globally convergent solution to the K-K-T system are obtained for both feasible and infeasible starting points under much weaker conditions.

1. Introduction

Throughout, let \( R^n \), \( R^n_+ \), and \( R^n_{++} \) denote the \( n \)-dimensional Euclidean space, nonnegative orthant, and positive orthant of \( R^n \), respectively. In the paper, the following general nonconvex nonlinear programming will be considered:

\[
\begin{align*}
\min & \quad f(x) \\
\text{s.t.} & \quad g_i(x) \leq 0, i = 1, 2, \ldots, m \\
& \quad h_j(x) = 0, j = 1, 2, \ldots, l,
\end{align*}
\]

where \( x \in R^n \) and \( f: R^n \rightarrow R \), \( g_i: R^n \rightarrow R^m \), and \( h_j: R^n \rightarrow R^l \) are three continuously differentiable functions. Denote \( \Omega = \{ x | g_i(x) \leq 0, i = 1, 2, \ldots, m, h_j(x) = 0, j = 1, 2, \ldots, l \} \) and \( \Omega^0 = \{ x | g_i(x) < 0, i = 1, 2, \ldots, m, h_j(x) = 0, j = 1, 2, \ldots, l \} \).

Homotopy method has been paid much attention as an important globally convergent computational method in finding solutions to various nonlinear problems since it was introduced and studied by Kellogg et al. [1], Smale [2], and Chow et al. [3]. However, the original homotopy is only single homotopy and needs much strong assumptions when solving nonlinear problems. In the 1990s, a combined homotopy interior point (CHIP) method was firstly proposed for solving nonconvex programming under the normal cone condition by Feng and Yu in [4]. From then on, various CHIP methods, as an efficiently implementable algorithm, were widely used and newly constructed for solving general nonconvex programming, fixed point problems, complementarity problems, variational inequality, and so on, see, e.g., [5–20].

In 2001, for reducing the dimension of the systems arising in the numerically tracing process and weakening convergent conditions, Yu et al. [21] proposed an aggregate constraint homotopy method (ACH method) for nonconvex programming by using the so-called aggregate function of the constraints. In 2018, a new ACH method for nonlinear programming problems with inequality and equality...
constraints was presented in [22]. However, the ACH method still belongs to CHIP since it requires the initial point which was also in the original feasible set. In 2006, to avoid the disadvantage of CHIP must choose the initial point in the feasible set, a constraint-shifting combined homotopy infeasible interior-point method in which the initial point can be chosen in both feasible and infeasible sets for solving nonlinear programming with only inequality constraints was proposed by Yu and Shang in [23, 24]. In 2012, to extend the constraint-shifting combined homotopy method to solve the general nonlinear programming, another new combined homotopy infeasible interior-point method for solving nonconvex programming with both inequality and equality constraints was proposed in [25], in which only inequality constraints need to satisfy the normal cone condition. From then on, more constraint-shifting homotopy equations were constructed and extended for solving nonlinear programming, principal-agent problem, fixed point problem, and so on, see, e.g., [26–30]. However, these combined homotopy methods usually required some cone conditions for proving the strong convergence of the existence of the smooth homotopy pathway.

By the enlightenment of the above references, without any cone condition, an aggregate constraint combined homotopy infeasible interior-point method for solving nonconvex nonlinear programming with both inequality and equality constraints is constructed, and the global convergence under much weaker conditions is obtained in the paper.

The remainder of this paper is organized as follows. In Section 2, the homotopy equation is constructed, and some lemmas from differential topology are introduced. In Section 3, the main results will be presented, and the existence and convergence of a smooth path from any given point in the infeasible set to the solution of KKT systems are proved. In Section 4, the numerical algorithm is presented.

2. Preliminaries

The following assumptions will be used:

(A1) $\Omega$ is a bounded and connected set, $\Omega^0 \neq \emptyset$.

(A2) $\forall x \in \Omega$, matrix $\{\nabla g_i(x)_{i \in \{1, \ldots, m\}}, \nabla h_j(x)\}$ is positive linearly independent at $x$, i.e.,

$$\sum_{i \in \{1, \ldots, m\}} \alpha_i \nabla g_i(x) + \sum_{j=1}^m \beta_j \nabla h_j(x) = 0, \quad \alpha_i \in \mathbb{R}, \beta_j \in \mathbb{R} \implies \alpha_i = \beta_j = 0.$$  

(2)

By [21], the aggregate function $g(x, \mu) = \mu \ln \left(\sum_{i=1}^m \exp(g_i(x)/\mu)\right)$, we have

$$\text{(I)} \nabla_x g(x, \mu) = \sum_{i=1}^m \tau_i(x, \mu)\nabla g_i(x), \quad \tau_i(x, \mu) \equiv \frac{\exp(g_i(x)/\mu)}{\sum_{j=1}^m \exp(g_j(x)/\mu)},$$  

(3)

$$\text{(II)} \ g(x) \leq g(x, \mu) \leq g(x) + \mu \ln m.$$  

We construct the following shifted aggregate constraint function only with inequality constraint functions:

$$\overline{g}(x, \mu) = \mu \phi(x) + (\theta - \mu)g(x, \mu),$$  

(4)

where $\theta \in (0, 1)$ is a parameter and $\phi(x)$ are convex and three continuously differentiable functions. Therefore, we have

$$\overline{g}(x, \theta) = \theta \phi(x)$$  

(5)

and

$$\lim_{\mu \to 0^+} \overline{g}(x, \mu) = \lim_{\mu \to 0^+} [\mu \phi(x) + (\theta - \mu)g(x, \mu)] = \theta g(x, \mu).$$  

(6)

Obviously, $\overline{g}(x, \mu)$ are also three continuously differentiable functions; let $\Omega^0_\mu = \{x|\overline{g}(x, \mu) \leq 0\}$, $\Omega^1_\mu = \{x|\overline{g}(x, \mu) < 0\}$, $\partial \Omega_\mu = \Omega^0_\mu \cup \Omega^1_\mu$, and $\Omega(0) = \Omega$.

Lemma 1 (see [21]). If assumptions (A1) and (A2) hold, then there exists $\theta_1 > 0$, for $\forall \mu \in (0, \theta_1)$, and

$$\nabla_x \overline{g}(x, \mu) \neq 0, \quad \forall x \in \partial \Omega(\mu).$$  

(7)

Lemma 2 (see [27]). If assumptions (A1) and (A2) and hold, there exists $\theta_2 > 0$, for $\forall \mu \in (0, \theta_2)$, and we have $\Omega_\mu$ which is a bounded and connected set, and $\Omega^0_\mu$ is nonempty.

Lemma 3. If assumptions (A1) and (A2) hold, there exists $\theta_3 > 0$, for any given feasible point $x$, and $\forall \mu \in (0, \theta_3)$, matrix $\{\nabla_x \overline{g}(x, \mu), \nabla h_j(x)\}$ is positively linearly independent.

Proof. Proved by contradiction. For $\forall \mu_k \in (0, \theta_3]$ and any feasible point $x^k$, there exists $\alpha^k \geq 0$ and $\beta^k_j \in \mathbb{R}$ belonging to real part, which are simultaneously not equal to zeros, such that

$$\alpha^k \nabla_x \overline{g}(x^k, \mu_k) + \sum_{j=1}^m \beta^k_j \nabla h_j(x^k) = 0.$$  

(8)

Let $\beta^k = \max_{1 \leq j \leq l} |\beta^k_j|$ and $\xi^k = \max\{\alpha^k, \beta^k\}$; divide both sides of (8) by $\xi^k$, i.e.,
\[ \frac{\alpha^k}{\xi^k} \nabla_x \tilde{g}(x^k, \mu_k) + \sum_{j=1}^{m} \frac{\beta_j^k}{\xi^k} \nabla h_j(x^k) = 0. \]  

(9)

When \( \mu_k \to 0, \ x_k \to x^*; \) let \((\alpha^k/\xi^k, \beta_j^k/\xi^k) \to (\alpha^*, \beta_j^*), \) as \( k \to \infty, \) and we have

\[ \lim_{k \to \infty} \left[ \frac{\alpha^k}{\xi^k} \nabla_x \tilde{g}(x^k, \mu_k) + \sum_{j=1}^{m} \frac{\beta_j^k}{\xi^k} \nabla h_j(x^k) \right] = 0. \]

(10)

Define \( \theta = \min[\theta_1, \theta_2, \theta_3]. \) Since \( \Omega_0^\alpha \) is nonempty, for any \( x^0 \in \Omega_0^\alpha, \ y^0 \in R^m, \) and \( z^0 \in R^l, \) let \( w^0 = (x^0, y^0, z^0), \) and we construct the homotopy equation as follows:

\[ H(w, w_0, \mu) = \begin{pmatrix} \nabla f(x) + \nabla x \tilde{g}(x, \mu \theta) y + \nabla h(x) z + \mu (x - x^0) \\ Y \tilde{g}(x, \mu \theta) + \mu \eta \\ h(x) - \mu z \end{pmatrix} = 0, \]

(11)

where \( w = (x, y, z) \in \Omega_\mu \times R^m \times R^l, \) \( \eta \in R^m, \) and \( Y = \text{diag}(y). \)

When \( \mu = 0, \) homotopy equation (11) turns to the K-K-T system

\[ \begin{pmatrix} \nabla f(x) + \theta \sum_{i=1}^{l} \nabla g_i(x) y_i + \sum_{j=1}^{l} \nabla h_j(x) z_j \\ Y \tilde{g}(x) \\ h(x) \end{pmatrix} = 0. \]

(12)

When \( \mu = 1, \) homotopy equation (11), \( H(w, w_0, 1) = 0, \) has a unique simple solution

\[ (x, y, z) = (x^0, y^0, z^0) = \left( x^0, -\left[ \text{diag}(\tilde{g}(x^0, \theta)) \right]^{-1} \eta, h(x^0) \right). \]

(13)

The following lemmas from differential topology will be used in the next section. At first, let \( U \subset R^n \) be an open set, and let \( \phi: U \to R^p \) be a \( C^\alpha \) mapping, where \( \alpha > \max\{0, n - k\}. \) If \( 0 \in R^k \) is a regular value of \( \phi, \) then for almost all \( a \in V, \) \( 0 \) is a regular value of \( \phi_a = F(a, \cdot). \)

**Lemma 4** (see [31]). Let \( V \subset R^n \) and \( U \subset R^m \) be open sets, and let \( \phi: V \times U \to R^p \) be a \( C^\alpha \) mapping, where \( \alpha > \max\{0, n - p\}. \) If \( 0 \in R^k \) is a regular value of \( \phi, \) then \( \phi^{-1}(0) \) consists of some \( (n - p) \)-dimensional \( C^\alpha \) manifolds.

**Lemma 5** (see [31]). Let \( \phi: U \subset R^n \to R^p \) be a \( C^\alpha \) mapping, where \( \alpha > \max\{0, n - p\}. \) If \( 0 \) is a regular value of \( \phi, \) then \( \phi^{-1}(0) \) consists of some \( (n - p) \)-dimensional \( C^\alpha \) manifolds.

**Lemma 6** (see [31]). A one-dimensional smooth manifold is diffeomorphic to a unit circle or a unit interval.

### 3. Main Results

For a given \( w^0 \in \Omega_0^\alpha \times R^m \times R^l, \) rewrite \( H(w, w_0, \mu) \) in homotopy equation (11) as

\[ H_{w_0}(w, \mu) = H(w, w_0, \mu). \]

(15)

The zero-point set of \( H_{w_0}(w, \mu) \) is

\[ H_{w_0}^{-1}(0) = \{(w, \mu) \in \Omega_0^\alpha \times R^m \times R^l \times (0, 1): H_{w_0}(w, \mu) = 0\}. \]

(16)
Lemma 7. If assumptions (A1) and (A2) hold, given \( w^0 \in \Omega(\theta) \times R^m \times R^l \), and if there exists a smooth curve \( \Gamma_{w^0} \) starting from \( (w^0, 1) \) in \( \Omega(\mu) \times R^m \times R^l \times (0, 1] \), then it must be bounded.

Proof. If \( \Gamma_{w^0} \in \Omega(\mu) \times R^m \times R^l \times (0, 1] \) is unbounded, there exists \( (x^k, y^k, z^k, \mu_k) \in \Gamma_{w^0} \), and \( \|y^k, z^k, \mu_k\|^2 \to \infty \).

From the second equation of (11), we have
\[
Y^k g(x^k, \mu_k \theta) = \mu_k \eta = 0. \tag{17}
\]

From the first equation of (11), we have
\[
(1 - \mu_k) \left( \nabla f(x^k) + \sum_{i=1}^m g_i(x^k, \mu_k \theta) f^i j + \nabla h_j(x^k) z^k_j \right) + \mu_k (x^k - x^0) = 0. \tag{19}
\]

(i) When \( \nu = 1 \), from the third equation of (11), we have \( z^k \to h(\bar{x}) \) as \( k \to \infty \), which implies that \( \{z^k\} \) is bounded. Hence, \( \|y^k\|^2 = \|z^k\|^2 \) is unbounded, and \( \bar{x} \in \Omega_{\mu} \). If \( \|1 - \mu_k\| \to 0 \), without loss of generality, and suppose
\[
\bar{x} = \lim_{k \to \infty} (1 - \mu_k) \left( \nabla f(x^k) + \sum_{i=1}^m g_i(x^k, \mu_k \theta) f^i j + \nabla h_j(x^k) z^k_j \right)
\]

but \( g(x, \theta) = \partial \phi(x) \) is a convex function; this is impossible.

If \( \|1 - \mu_k\| y^k \to \infty \), the discussion is the same as the following case (ii).

(ii) When \( \nu \in (0, 1) \), without loss of generality, suppose that \( (1 - \mu_k) y^k / \|1 - \mu_k\| y^k \to \bar{\alpha} \) with \( \|\bar{\alpha}\| = 1 \) and \( \bar{\alpha}_i = 0 \) for \( i \notin I_{\nu}(\bar{x}) \). Through dividing both sides of equation (19) by \( \|1 - \mu_k\| y^k \) and taking limits, we have
\[
\sum_{i \notin I_{\nu}(\bar{x})} \bar{\alpha}_i \nabla x g_i (\bar{x}, \bar{\nu} \theta) = 0,
\]

which contradicts with Lemma 3.

By equation (17), \( Y^k g(x^k, \mu_k \theta) \leq 0, i = 1, 2, \ldots, m \), \( x^k \in \Omega(\mu_0) \) and \( \mu_k \theta \leq \theta \), and by Lemma 2, \( \Omega_{\mu_0} \) is also bounded, so \( \{x^k\} \) is a bounded sequence. Therefore, \( \{x^k\} \) must exist a convergent subsequence which is also denoted as \( \{x^k\} \). Let \( x^k \to \bar{x} \) and \( \|y^k, z^k\| \to \infty \) as \( k \to \infty \). Denoting \( I^* = \{i \in [1, 2, \ldots, m] | y^k \to \infty \} \), by (17), \( I^* \subset \pi_{\mu} (\bar{x}) \); therefore, we obtain
\[
(1 - \mu_k) y^k \to \bar{y}, \text{ then } \bar{y}_i = 0 \text{ for } i \notin I^* \text{ from the second equation of (11). Taking limits in (19), we have}
\]

(iii) When \( \nu = 0 \), without loss of generality, suppose that \( (y^k, z^k) / \|y^k, z^k\| \to (\bar{\alpha}, \bar{t} \bar{\beta}) \) with \( \|\bar{\alpha}, \bar{t} \bar{\beta}\| = 1 \) and \( \alpha_i = 0 \) for \( i \notin I_0 (\bar{x}) \). Through dividing both sides of equation (19) by \( \|y^k, z^k\| \) and taking limits, we have
\[
\sum_{i \notin I_0(\bar{x})} \bar{\alpha}_i \nabla x g_i (\bar{x}, 0) + \sum_{j=1}^l \bar{\beta}_j \nabla h_j (\bar{x}) = 0,
\]

which contradicts with Lemma 3.

In conclusion, from the above discussion, we obtain that \( \Gamma_{w^0} \) is a bounded curve in \( \Omega_{\mu} \times R^m \times R^l \times (0, 1] \). \( \square \)

Theorem 1. Suppose assumptions (A1) and (A2) hold, for almost all \( w^0 \in \Omega_0 \times R^m \times R^l \), the zero-point set of homotopy equation (11) contains a smooth curve \( \Gamma_{w^0} \subset \Omega_0 \times R^m \times R^l \times (0, 1] \).
(0, 1] starting from \((x^0, y^0, z^0, 1)\), which terminates or approaches to the hyperplane \(\mu \to 0\). If \((x^*, y^*, z^*, 0)\) is a limit point of \(\Gamma_w^0\), then \(w^* = (x^*, y^*, z^*)\) is a solution to the K-K-T system of problem (11).

**Proof.** Let \(\tilde{H}(w, x^0, \mu): \Omega_{w^0} \times R_m^+ \times R^l \times \Omega_{x^0} \times \Omega_{\mu^0} \times (0, 1] \to \Omega_{w^0} \times R_m^+ \times R^l\) be the same map as \(H(w, x^0, \mu)\) but taking \(x^0\) as variate. Consider the following submatrix of the Jacobian

\[
\begin{pmatrix}
\ast & \ast & -\mu I \\
\end{pmatrix}
\]

\[
\frac{\partial \tilde{H}(w, x^0, \mu)}{\partial (y, z, x^0)} = \text{diag} \left( \tilde{g}(x, \mu \theta) \right)
\begin{pmatrix}
0 & 0 \\
0 & 0 \\
\end{pmatrix}
\]

(23)

For all \(\mu \in (0, 1]\) and any \(x^0 \in \Omega_{x^0}\) from \((w, \mu) \in H^{-1}\) and \(Y \tilde{g}(x, \mu \theta) + \mu \eta = 0\), we get that \(\text{diag} \left( \tilde{g}(x, \mu \theta) \right)\) is nonsingular, which implies that \(\partial \tilde{H}(w, x^0, \mu) / \partial (y, z, x^0)\) is nonsingular.

Hence, matrix \(D \tilde{H}(w, x^0, \mu)\) is full row rank. That is, 0 is a regular value of \(\tilde{H}(w, x^0, \mu)\). By Lemma 4, we have that, for almost all \(x^0 \in \Omega_{x^0}^0, 0\) is a regular value of \(H(w, x^0, \mu)\).

Note that the matrix

\[
\frac{\partial \tilde{H}(w^0, x^0, 1)}{\partial w} = \begin{pmatrix}
I & 0 & 0 \\
0 & \text{diag} \left( \tilde{g}(x^0, \theta) \right) & 0 \\
0 & 0 & -I
\end{pmatrix}
\]

(24)

is nonsingular. From Lemma 5, if 0 is a regular value of \(H(w, w^0, \mu), \partial \tilde{H}(w, w^0, 1) / \partial w\) is nonsingular, and the fact \(H(w^0, w^0, 1) = 0, H^{-1}_w(0)\) must contain a smooth curve \(\Gamma_{w^0}\) starting from \((x^0, y^0, z^0, 1)\) and going to \(\Omega_{w^0} \times R_m^+ \times R^l \times [0, 1]\). Then, from Lemma 6, the curve \(\Gamma_{w^0} \subset \Omega_{w^0} \times R_m^+ \times R^l \times (0, 1]\) must be diffeomorphic to a unit circle or a unit interval \([0, 1]\).

We have that \(\Gamma_{w^0}\) is not diffeomorphic to a unit circle. That is, \(\Gamma_{w^0}\) is diffeomorphic to \([0, 1]\). Let \((x^*, y^*, z^*, \mu_s)\) be a limit point of \(\Gamma_{w^0}\); only the following five cases are possible:

(i) \((x^*, y^*, z^*) \in \Omega_{\mu_s} \times R_m^+ \times R^l \mu_s = 1 \| (y^*, z^*) \| < \infty, \)

(ii) \((x^*, y^*, z^*) \in \Omega_{\mu_s} \times R_m^+ \times R^l \mu_s = [0, 1] \| (y^*, z^*) \| = \infty, \)

(iii) \((x^*, y^*, z^*) \in \Omega_{\mu_s} \times \partial R_m^+ \times R^l \mu_s = (0, 1] \| (y^*, z^*) \| < \infty, \)

(iv) \((x^*, y^*, z^*) \in \partial \Omega_{\mu_s} \times \partial R_m^+ \times R^l \mu_s = (0, 1] \| (y^*, z^*) \| < \infty, \)

(v) \((x^*, y^*, z^*) \in \Omega \times R_m^+ \times R^l \mu_s = 0 \| (y^*, z^*) \| < \infty. \)

By differentiating (26), we can get

\[
H'(w(s), \mu(s)) \begin{pmatrix}
w \\
\mu
\end{pmatrix} = 0,
\]

(27)

\[
\mu(0) = w^0, \mu(0) = 1,
\]

where \(H'\) is the derivative of \(H\).

As how to trace the homotopy path \(\Gamma_{w^0}\) numerically, we can use the standard predictor-corrector procedure; for more details, see [32, 33]. In this paper, our contribution is only the theoretical results about the proposed algorithm which only requires that any initial point can be chosen in the shifted feasible set but not necessarily in the original feasible set. The relative homotopy algorithms and numerical simulations on the performance for the proposed algorithm can be implemented as [34].

**Data Availability**

No data were used to support this study.

**Conflicts of Interest**

The authors declare no conflicts of interest.
Authors’ Contributions
Both authors contributed equally to this paper and read and approved the final manuscript.

Acknowledgments
Zhichuan Zhu was supported by the Education Department Foundation of Jilin Province (Grant no. JJKH20190742SK) and the Advanced Talents Research Fund of Liaoning University. Yeong-Cheng Liou was partially supported by MOST 109-2410-H-037-010 and MOST 108-2410-H-037-020 and the Kaohsiung Medical University Research Foundation (Grant no. KMU-M108002).

References


