Combined Homotopy Method for Inequality Constrained Problems

Qingqun Huang
School of Mathematics and Statistics, Hechi University, Yizhou Guangxi
Email: 44584330@qq.com

Received: May 4th, 2016; accepted: May 23rd, 2016; published: May 26th, 2016

Abstract
For the optimization problem with inequality constraints, this paper constructs a new homotopy equation which with the Newton's method to get a combined homotopy Newton algorithm. The global linear convergence of the algorithm is proved at the end.

Keywords
Combined Homotopy, Convex Nonlinear Programming, Global Convergence, Newton Method

组合同伦法求不等式约束问题

黄青群
河池学院数学与统计学院，广西宜州
Email: 44584330@qq.com

收稿日期: 2016年5月4日；录用日期: 2016年5月23日；发布日期: 2016年5月26日

摘　要
对含不等式约束的优化问题，构造一个新的同伦方程，与牛顿法相结合得到一个组合同伦牛顿算法，最后给出了该算法的全局线性收敛性的证明。

http://dx.doi.org/10.12677/orf.2016.62008
1. 引言


同伦方法的基本思想是：为求解非线性方程组

\[ F(x) = 0, \]

其中 \( F: \Omega \subset \mathbb{R}^n \to \mathbb{R}^n \) 是光滑映射，构造带参数 \( t \) 的同伦映射 \( H(x,t) : \Omega \times (0,1] \to \mathbb{R}^n \)，使之满足

\[ H(x,1) = G(x), H(x,0) = F(x), \]

其中 \( G(x) = 0 \) 的解 \( x^0 \) 为已知。只要 \( H \) 构造得合适，在特定的条件下，同伦方程就可确定一条从 \( (x^0,1) \) 出发，趋于超平面 \( t = 0 \) 的光滑曲线，称之为同伦路径，而该曲线另一端的任意极限点的 \( x \) 分量 \( x^* \) 就是 \( F(x) = 0 \) 在 \( \Omega \) 中的解。因此可通过数值跟踪该曲线从而得到 \( F(x) = 0 \) 的解。本文对含不等式约束的优化问题提出了一个同伦路径跟踪方法，即通过构造一个新的同伦方程，与牛顿法相结合的得到一个组合同伦牛顿算法，最后给出了该算法的全局线性收敛性的证明。

考虑下列不等式约束规划问题:

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

其中 \( x \in \mathbb{R}^n \)，\( f(x) \)，\( g_i(x) \)，\( i = 1, 2, \ldots, m \) 为充分光滑函数，且 \( f(x) \) 为严格凸函数，\( g_i(x) \)，\( i = 1, 2, \ldots, m \) 为凸函数。定义问题(1)的可行域集合为 \( \Omega = \{ x \in \mathbb{R}^n : g_i(x) \leq 0, \quad i = 1, 2, \ldots, m \} \)，严格可行域集合为 \( \Omega^0 = \{ x \in \mathbb{R}^n : g_i(x) < 0, \quad i = 1, 2, \ldots, m \} \)。如果 \( \Omega \neq \emptyset \) ，则问题(1)的解存在的充要条件是存在 \( x,y \) 为下列K-K-T方程的解:

\[ H(\omega) = \left( \nabla f(x) + \nabla g(x)y \right) \begin{bmatrix} y \\ Y \end{bmatrix} = 0, \quad g(x) \leq 0, \quad y \geq 0 \tag{2} \]

其中

\[ \omega = (x,y)^T, \quad g = (g_1, \ldots, g_m)^T, \quad \nabla g(x) = (\nabla g_1(x), \ldots, \nabla g_m(x))^T, \quad Y = \text{diag}(y), \quad \nabla f(x) = \left( \frac{\partial f(x)}{\partial x} \right)^T, \]

\( y \in \mathbb{R}^m \) 为是(1)的约束拉格朗日乘子。

为了求解问题(2)，构造如下同伦方程:

\[ H(\omega, \mu) = \left( 1 - \mu \right) \left( \nabla f(x) + \nabla g(x)y \right) + \mu \left( \epsilon - e^{(0)} \right) \begin{bmatrix} y \\ Y \end{bmatrix} = 0 \tag{3} \]

其中 \( \mu \in (0,1], \quad x^{(0)} \in \Omega^0, \quad \epsilon = (1, \ldots, 1)^T \)。我们作如下的假设:

(C1) \( \Omega^0 \) 非空有界。
(C2) \( \forall x \in \partial \Omega \), \( \{ V g_i(x) : i \in B(x) \} \) 是列满秩矩阵，其中 \( B(x) = \{ g_i(x) = 0, i = 1, 2, \ldots, m \} \)。

(C3) (\( \Omega \) 在 \( x \) 处的法锥条件) \( \forall x \in \partial \Omega \), \( \Omega \) 在 \( x \) 点的法锥与 \( \Omega \) 仅交于点 \( x \)。即 \( \forall x \in \partial \Omega \)，有

\[
\left\{ x + \sum_{i \in I} V g_i(x) y_i : y_i \geq 0, i \in B(x) \right\} \cap \Omega = \{ x \}.
\]

引理1 \( f(x), g_i(x), i = 1, 2, \ldots, m \) 为充分光滑函数，\( \omega^{(0)} \in \Omega^0 \times R_m^+, \) 同伦方程(3)定义，则对任意的 \( \mu \in (0, 1), (\omega, \mu) \in \Omega^0 \times R_m^+ \times (0, 1) \)，\( H'_\omega(\omega, \mu) \) 是非奇异的。其中 \( R_m^+ = \{ x \in R^m : x > 0 \} \)。

证明：通过计算整理，有

\[
H'_\omega(\omega, \mu)^T = \begin{pmatrix}
\frac{\partial H(\omega, \mu)}{\partial \omega}
\end{pmatrix}^T = \begin{pmatrix}
(1-\mu)Q(\omega) + \mu e^T (1-\mu)\n∇g(x)^T G(x)
\end{pmatrix}.
\]

其中 \( Q(\omega) = \nabla^2 f(x) + \sum_{i \in I} \nabla^2 g_i(x), I = \{ 1, 2, \ldots, m \} \)，而 \( G(x) = \text{diag}(g(x)) \) 为负对角矩阵，经过矩阵初等变换后，可得 \( H'_\omega(\omega, \mu)^T \) 的相似阵

\[
\begin{pmatrix}
(1-\mu)Q(\omega) + \mu e^T + W(\omega) & 0 \\
∇g(x)^T & G(x)
\end{pmatrix}
\]

其中

\[
W(\omega) = -(1-\mu)\left(∇g(x)^T\right)^T G(x)^{-1} ∇g(x)^T.
\]

有

\[
\| H'_\omega(\omega, \mu) \| = \| G(x) \| \cdot \left(1-\mu\right)\| Q(\omega) + \mu e^T + W(\omega) \|,
\]

因为 \( f(x), g_i(x), i \in I \) 是充分光滑的凸函数，所以 \( \nabla^2 f(x), \nabla^2 g_i(x) \) 为半正定矩阵，又 \( \chi_i > 0, i \in I \)，从而 \( Q(\omega) \) 为半正定矩阵，因 \( \chi \) 为正对角矩阵，通过计算知 \( W(\omega) \) 也为半正定矩阵，所以对任意的 \( \mu \in (0, 1), (\omega, \mu) \in \Omega^0 \times R_m^+ \times (0, 1) \)，有 \( 1-\mu)Q(\omega) + \mu e^T + W(\omega) \) 为正定矩阵，\( |G(x)| = 0 \)。所以 \( H'_\omega(\omega, \mu) \neq 0 \)，即 \( H'_\omega(\omega, \mu) \) 是非奇异矩阵。

根据引理1 及隐函数定理，对任意的 \( \mu \in (0, 1), \) 同伦方程(3)都有唯一的一个 \( \omega(\mu) \) 与之对应。由点集 \{ \omega(\mu): \mu \in (0, 1) \} 构成 \( \Omega \times R_m^+ \) 内的连续函数，即为组合同伦曲线。\( \mu = 1 \) 时 \( \omega(1) = \omega^{(0)} \)，\( \mu \rightarrow 0 \) 时 \( \omega(\mu) \rightarrow \omega^* = (x^*, y^*) \)，\( x^* \) 即为问题(1)的最优解。问题(1)的解可跟踪这条曲线得到。下面第2节构造的组合同伦牛顿算法就是跟踪这条曲线从而得到一系列迭代点：\{ \omega(\mu): \mu \in (0, 1) \} \subset \Omega^0 \times R_m^+，且问题(1)的解就是其极限点 \( x^* \)。第3节讨论了算法的收敛性。

2. 算法

设组合同伦方程(3)的解曲线为 \( \Gamma \)，且当 \( \mu = 1 \) 时 \( \Gamma \) 关于 \( \omega^{(0)} \) 的分量是 \( \omega^{(0)} \)。由文献[8, 9]知当 \( \mu \rightarrow 0 \) 时，\( \Gamma \) 中关于 \( \omega^{(0)} \) 极限点 \( \omega^* \) 存在等价于问题(1)的 K-K-T 方程(2)有解 \( \omega^* \)。

定义1：令

\[
N(\beta, \mu) = \{ (\omega, \mu): H(\omega, \mu) \leq \beta \mu, \mu \in (0, 1) \},
\]

称 \( N(\beta, \mu) \) 为 \( \beta \)-锥邻域，其中 \( \beta > 0 \) 为邻域半径。
引理 2：设 \( y > 0, \quad \beta \in (0,1) \)，那么 \( N(\beta, \mu) \subset \Omega^v \times R^v_+ \times (0,1] \)。

证：由于任意的 \((\omega, \mu) \in N(\beta, \mu)\)，有 \( |H(\omega, \mu)| \leq \beta \mu \)，所以有 \(|H(\omega, \mu)| \leq \beta \mu \)。再由方程(3)的第二个等式有 \( |y g(x) + \mu e| \leq \beta \mu \Rightarrow -\beta \mu \leq y g(x) + \mu \leq \beta \mu \Rightarrow y g(x) \leq (\beta - 1) \mu < 0 \)。

由已知 \( y > 0 \)，因此有 \( g(x) < 0 \)。综上所述有 \( N(\beta, \mu) \subset \Omega^v \times R^v_+ \times (0,1] \)。

组合同伦牛顿算法

**Step 0 (初始化)**
令 \( k = 0 \) ， \( \mu_0 = 1 \) ， \( \beta_0 = 1 \) ， \( \beta \in (0, \beta_0) \)， \( (\omega^0, 1) \in N(\beta, 1) \) ， \( \alpha = (0,1) \) ， \( \delta \in (0,1) \) ， \( \varepsilon > 0 \)。

**Step 1 (终止条件)**
如果 \( \mu_k < \varepsilon \) ，则算法停止，且 \( \omega^k = (x^k, y^k)^T \) 即为方程(3)的近似解。

**Step 2 (计算牛顿方向)**

\[
H'(\omega^k, \mu_k) \Delta \omega^k = -H(\omega^k, \mu_k)
\] (4)

其中 \( \Delta \omega^k = (\Delta x^k, \Delta y^k) \)。

**Step 3 (线搜索)**
取 \( \lambda_k \) 为 \( 1, \delta, \delta^2, \cdots \) 中的最大值，使之满足

\[
\|H(\omega^k + \lambda_k \Delta \omega^k, (1-\alpha \lambda_k) \mu_k)\| \leq (1-\alpha \lambda_k) \beta \mu_k
\] (5)

令 \( \omega^{k+1} = \omega^k + \lambda_k \Delta \omega^k \) ， \( \mu_{k+1} = (1-\alpha \lambda_k) \mu_k \)， \( k := k + 1 \) ，返回 Step 1。

3. 算法的全局收敛性

先讨论函数 \( H(\omega, \mu) \) 以及算法的一些性质。

**引理 3 [4]**：假设 \( H(\omega, \mu) \) 是由方程(3)定义的，给定有界凸集 \( M \subset R^m \) ，则对任意的 \( \omega \in M \) ， \( \mu \in (0,1] \)，存在常数 \( c_1, c_2 > 0 \) ，使得

\[
\|H'(\omega, \mu)\| \leq c_1 , \quad \|H^{-1}(\omega, \mu)\| \leq c_2 .
\]

**定理 1**：算法是良定的。

证明：由引理 1 知，算法的 Step 2 良定。又由文献[4]知，当

\[
\lambda_k \leq \frac{\beta - c_2 \alpha - \alpha \beta}{c_1 \beta^2 - c_2 \alpha}
\]

时，算法的 Step 3 良定，其中 \( c_1, c_2 > 0 \) 为常数。所以本文算法良定。

下面给出算法的全局收敛性。

**定理 2**：设 \( \{A(\omega^k, \mu_k)\} \) 为由算法产生的无穷序列，则

(i) 对 \( k = 0, 1, 2, \cdots \)

\[
(\omega^k, \mu_k) \in N(\beta, \mu_k)
\] (6)

\[
(1-\alpha \lambda_{k-1}) \cdots (1-\alpha \lambda_0) = \mu_k
\] (7)

(ii) 序列 \( \|H(\omega^k)\| \) 及 \{\mu_k\} 全局收敛于 0。

(iii) 存在 \( \omega^* = (x^*, y^*) \)，使得 \( \omega^{(k)} \to \omega^* , k \to \infty \)。那么 \( \omega^* \) 为同伦方程(3)的最优解， \( x^* \) 为问题(1)的最
优解。

证明：
(i) 对 \( k \) 进行数学归纳法证明。
当 \( k = 0 \) 时，由算法知 \( (\omega^0, \mu_0) = (\omega^0, 1) \in N(\beta, 1) \)，\( \mu_0 = 1 \)，所以结论成立。假定对任 \( k > 0 \) 有结论成立。那么对 \( k + 1 \)，由算法可得到 \( (\omega^{k+1}, \mu_{k+1}) \in N(\beta, \mu_{k+1}) \)，其中 \( 0 < \mu_{k+1} < \mu_k \)，且
\[
\mu_{k+1} = (1 - \alpha \lambda_k) \mu_k = (1 - \alpha \lambda_k)(1 - \alpha \lambda_{k-1}) \cdots (1 - \alpha) \lambda_0,
\]
所以(6)式成立。
(ii) 先证明 \( \{ \mu_k \} \) 全局收敛于 0。
由定理 1 可知，对任意的 \( k \geq 0 \)，有
\[
\frac{(\beta - c_2 \alpha - c_2 \beta)}{c_2 \beta^2 - c_2 \alpha} \leq 0.
\]
结合(7)式可得
\[
\mu_k \leq (1 - \alpha \lambda_k) \mu_0 \to 0 (k \to \infty).
\]
故有 \( \{ \mu_k \} \) 全局收敛于 0。
(2) 再证明 \( \| H(\omega^k) \| \) 全局收敛于 0。
因为
\[
H(\omega^k, \mu_k) - H(\omega^k) = \mu_k \left( \left( - \frac{\nabla f(x^k) + \nabla g(x^k)x^k + x^k - x^k}{e} \right) \right).
\]
所以必存在常数 \( c_5 > 0 \)，使得 \( \| H(\omega^k, \mu_k) - H(\omega^k) \| \leq c_5 \mu_k \)。根据前面所证，可得
\[
\| H(\omega^k) \| \leq \| H(\omega^k, \mu_k) \| + \| H(\omega^k, \mu_k) - H(\omega^k) \| \leq (\beta + c_5) \mu_k \to 0 (k \to \infty)
\]
所以序列 \( \{ H(\omega^k) \} \) 全局收敛于 0。
(iii) 因为
\[
\| \Delta \omega^k \| = \| H^{-1}(\omega^k, \mu_k) H(\omega^k, \mu_k) \| \leq c_5 \| H(\omega^k, \mu_k) \| \leq c_5 \beta \mu_k \to 0 (k \to \infty),
\]
因此 \( \{ \omega^k \} \) 柯西收敛于某点 \( \omega^* = (x^*, y^*) \)，又从 \( (\omega^k, \mu_k) \in N(\beta, \mu_k) \) 可知 \( \omega^* \in \Omega^0 \times \Omega^* \)。所以，\( \omega^* \) 是同伦方程(3)的最优解，因而 \( x^* \) 是问题(1)的最优解。

基金项目

广西高校科研项目(2013LX120)，河池学院教改课题(2014EB019)。

参考文献 (References)

Problems.

Applied Mathematics and Computation, 80, 209-224.
http://dx.doi.org/10.1016/0096-3003(95)00295-2

Applied Mathematics and Computation, 84, 193-211.
http://dx.doi.org/10.1016/S0096-3003(96)00086-0