

非線形最適化問題に対する主双対内点法の収束性について

矢部 博(東京理科大学理学部応用数学科)

1. はじめに

本講演では、非線形最適化問題

$$\begin{aligned} & \text{minimize} && f(x), \quad x \in \mathbf{R}^n, \\ & \text{subject to} && g(x) = 0, \quad x_i \geq 0, \quad i \in I_P \subset \{1, 2, \dots, n\} \end{aligned}$$

を解くための主双対内点法を考える。ただし、 $f: \mathbf{R}^n \rightarrow \mathbf{R}$, $g: \mathbf{R}^n \rightarrow \mathbf{R}^m$, I_P は添字集合で $p = |I_P|$ とする。ここで、行ベクトルが $e_i^t, i \in I_P$ から成る $p \times n$ 行列を E とすれば、非負条件は $x' \equiv Ex \geq 0$ と書ける。Lagrange 関数を

$$L(w) = f(x) - y^t g(x) - z^t Ex$$

と定義したとき、上記の最適化問題に対する Karush-Kuhn-Tucker (KKT) 条件は

$$(1) \quad r_0(w) = \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ X'Ze \end{pmatrix} = 0, \quad x' \geq 0, z \geq 0,$$

で表される。ただし、 $w = (x, y, z)^t$ とし、 $y \in \mathbf{R}^m$ と $z \in \mathbf{R}^p$ はそれぞれ等式制約条件と非負条件に対応する Lagrange 乗数であり、また、 $A(x)$ をその行ベクトルが $\nabla g_i(x)^t$ からなる $m \times n$ 行列、 $X' = \text{diag}(x'_1, \dots, x'_p)$, $Z = \text{diag}(z_1, \dots, z_p)$, $e = (1, \dots, 1)^t \in \mathbf{R}^p$ とする。

主双対内点法の基本的な考えは、バリアパラメータ μ を導入して相補条件 ($XZe = 0$) を $XZe = \mu e$ に置き換えて得られるバリア KKT 条件にニュートン法を適用するものである [10]。その際、直線探索法の枠組みで大域的収束性を実現するために、Yamashita [12] が提案した l_1 型バリアペナルティ関数が主変数 x に対するメリット関数としてよく使われているが、 (y, z) 空間では別々にステップ幅を選ぶ必要がある。そこで本講演では、主双対空間 $((x, z)$ 空間) におけるメリット関数を構築することによって、ステップ幅の選び方を統一することを考える。提案するメリット関数は l_1 型バリアペナルティ関数を利用したものである。そして提案する主双対内点法の大域的収束性を示す。最後に、局所的な収束速度についても触れる。

なお、直線探索法を用いた場合の大域的収束性については [1], [2], [12] 等を、信頼領域法を用いた場合の大域的収束性については [3], [4], [6], [14] 等を参照されたい。 l_2 型ペナルティ関数に基づいたメリット関数を用いた場合のアプローチが [8], [15] 等で研究されている。さらに、局所的な収束性 (収束速度など) に関する研究は [5], [7], [9], [11], [13] 等を参照されたい。

以下では、添字 k はアルゴリズムの外部反復や内部反復の反復回数を表し、 $\|\cdot\|$ は l_2 ベクトルノルムおよびそれから導かれる作用素ノルムを表す。また $S = \{x \in \mathbf{R}^n \mid x' > 0\}$, $\mathbf{R}_+^p = \{z \in \mathbf{R}^p \mid z > 0\}$ とおく。

2. バリア KKT 条件と主双対メリット関数

$$F_B(x; \mu) = f(x) - \mu \sum_{i=1}^p \log x'_i, \quad x \in S$$

で定義されるバリヤ関数を考える．ただし， $\mu > 0$ はバリヤパラメータである．このとき，もとの最適化問題の代わりに次の問題

$$(2) \quad \begin{array}{ll} \text{minimize} & F_B(x; \mu), \quad x \in S \\ \text{subject to} & g(x) = 0 \end{array}$$

を考えると，この問題に対する最適性の 1 次の必要条件は

$$\nabla f(x) - A(x)^T y - \mu E^t (X')^{-1} e = 0, \quad g(x) = 0, \quad (x' > 0)$$

で与えられる．ここで $z = \mu (X')^{-1} e \in \mathbf{R}^p$ とおけば， $\nabla f(x) - A(x)^T y - z = \nabla_x L(w)$ ， $XZe = \mu e$ となるので，上記の最適性条件は

$$(3) \quad r(w; \mu) \equiv \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ XZe - \mu e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (x > 0, \quad z > 0)$$

で表される．これをバリヤ KKT 条件という．これはちょうど KKT 条件 (1) で，相補条件 $XZe = 0$ を $XZe = \mu e$ で置き換えたものである．

他方，問題 (2) に対して次の l_1 型バリヤペナルティ関数

$$F_0(x; \mu) = f(x) - \mu \sum_{i=1}^p \log x'_i + \rho \sum_{i=1}^m |g_i(x)|$$

が提案されている [12]．ただし， $\rho > 0$ はペナルティパラメータである．この l_1 型バリヤペナルティ関数を実制約最小化する問題の最適性の 1 次の必要条件は，ペナルティパラメータ ρ が十分に大きい場合にはバリヤ KKT 条件 (3) で表されることが知られている．

バリヤ KKT 条件にニュートン法を適用すれば，ニュートン方向 $\Delta w = (\Delta x, \Delta y, \Delta z)^t$ は次の連立 1 次方程式 (ニュートン方程式) の解として定義される．

$$J(w) \Delta w = -r(w, \mu), \quad J(w) = \begin{pmatrix} G & -A(x)^t & -E^t \\ A(x) & 0 & 0 \\ ZE & 0 & X' \end{pmatrix},$$

ただし，行列 G はヘッセ行列 $\nabla_x^2 L(w)$ もしくはその準ニュートン近似である．

バリヤペナルティ関数 $F_0(x + s; \mu)$ の 1 次近似を

$$F_l(x; s) \equiv f(x) + \nabla f(x)^T s - \mu \sum_{i=1}^p \left(\log x'_i + \frac{(Es)_i}{x'_i} \right) + \rho \sum_{i=1}^m |g_i(x) + \nabla g_i(x)^T s|$$

と定義したとき，任意方向 s に対する関数 F_0 とその 1 次近似との差 ΔF_l は

$$\begin{aligned} \Delta F_l(x; s) &\equiv F_l(x; s) - F_0(x; \mu) \\ &= \nabla f(x)^T s - \mu e^T (X')^{-1} Es + \rho \sum_{i=1}^m |g_i(x) + \nabla g_i(x)^T s| - \rho \sum_{i=1}^m |g_i(x)| \end{aligned}$$

で与えられる．これは方向微係数に相当するものである．

次に直線探索法で用いるメリット関数について考える．我々は，次の l_1 型主双対メリット関数を提案する．ただし $\rho' > 0$ である．

$$F(x, z; \mu) = F_0(x; \mu) + \rho' \phi(x, z), \quad \phi(x, z) = \log \frac{(x')^t z / p + \|X'z - \mu e\|^2}{\left(\prod_{i=1}^p x'_i z_i \right)^{1/p}}$$

3. アルゴリズムと大域的収束性

バリエーション KKT 条件を用いた主双対内点法のアルゴリズムは以下の通りである。

[アルゴリズム IP](外部反復)

Step 0. (初期設定) パラメータ $\varepsilon > 0$, $M_c > 0$ および, 正の数列 $\{\mu_k\}$, $\mu_k \downarrow 0$ を与える. $k = 0$ とおく.

Step 1. もし $\|r_0(w_k)\| \leq \varepsilon$ ならば停止する.

Step 2. 近似バリエーション KKT 条件 $\|r(w_{k+1}, \mu_k)\| \leq M_c \mu_k$ を満たす点 w_{k+1} を求める.

Step 3. $k := k + 1$ において Step 1 へ行く. □

このアルゴリズムの Step 2 (近似バリエーション KKT 点の計算) の部分は, 次のアルゴリズムとして記述出来る.

[アルゴリズム LS](内部反復)

バリエーションパラメータ $\mu > 0$ と正定数 M_c が与えられる.

Step 0. (初期設定) $w_0 \in S \times \mathbf{R}^m \times \mathbf{R}_+^p$, $\rho > 0$, $\gamma \in (0, 1)$, $\beta \in (0, 1)$, $\varepsilon_0 \in (0, 1)$ を与える. $k = 0$ とおく.

Step 1. もし $\|r(w_k, \mu)\| \leq M_c \mu$ ならばアルゴリズム IP(外部反復) に戻る.

Step 2. ニュートン方程式 $J(w_k) \Delta w_k = -r(w_k, \mu)$ の解 Δw_k を探索方向に選ぶ. ただし

$$J(w_k) = \begin{pmatrix} G_k & -A(x_k)^t & -E^t \\ A(x_k) & 0 & 0 \\ Z_k^t E & 0 & X_k' \end{pmatrix},$$

G_k はヘッセ行列 $\nabla_x^2 L(w_k)$ もしくはその近似行列である.

Step 3. $\bar{\alpha}_k = \min \{\gamma \alpha_{k\max}, 1\}$,

$$\alpha_{k\max} = \min \left\{ \min_i \left\{ -\frac{(x_k')_i}{(\Delta x_k')_i} \mid (\Delta x_k')_i < 0 \right\}, \min_i \left\{ -\frac{(z_k)_i}{(\Delta z_k)_i} \mid (\Delta z_k)_i < 0 \right\} \right\}$$

として, 次のアルミホ基準

$$F(x_k + \bar{\alpha}_k \beta^{l_k} \Delta x_k, z_k + \bar{\alpha}_k \beta^{l_k} \Delta z_k; \mu) - F(x_k, z_k; \mu) \leq \varepsilon_0 \bar{\alpha}_k \beta^{l_k} \Delta F(x_k, z_k; \Delta x_k, \Delta z_k)$$

を満足する最小の非負整数 l_k を求めて $\alpha_k = \bar{\alpha}_k \beta^{l_k}$ とおく. ただし,

$$\Delta F(x, z; \Delta x, \Delta z) = \Delta F_l(x; \Delta x) + \rho' \nabla \phi(x, z)^t \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix}.$$

Step 4. $w_{k+1} = w_k + \alpha_k \Delta w_k$, $k := k + 1$ において Step 1 へ行く. □

大域的収束性示すために次の条件を仮定する.

(G1) f と g は 2 回連続的に微分可能である.

(G2) 初期点における $F(x, z; \mu)$ の準位集合 $\{(x, z) \in S \times \mathbf{R}_+^p \mid F(x, z; \mu) \leq F(x_0, z_0; \mu)\}$ はコンパクトである.

(G3) (G2) で定義された準位集合において $\text{rank } A(x) = m$ が成り立つ.

(G4) 任意の k に対して $\rho > \|y_k + \Delta y_k\|_\infty$ が成り立つ.

(G5) 行列 G_k は一様に正定値かつ有界である. □

このとき, 次の定理によって探索方向が $F(x, z; \mu)$ の降下方向になることが評価される.

[定理 1] 探索方向 $(\Delta x_k, \Delta z_k)$ は次式を満たす .

$$\begin{aligned} \Delta F(x_k, z_k; \Delta x_k, \Delta z_k) \leq & -\Delta x_k^t (G_k + E^t (X_k')^{-1} Z_k E) \Delta x_k - (\rho - \|y_k + \Delta y_k\|_\infty) \|g(x_k)\|_1 \\ & - \rho' \frac{\|X_k' Z_k e - \mu e\|^2}{(x_k')^t z_k / p + \|X_k' z_k - \mu e\|^2} \end{aligned}$$

この定理より , 固定された $\mu > 0$ に対するアルゴリズム LS の大域的収束性が示せる .

[定理 2] アルゴリズム LS によって生成される点列 $\{w_k\}$ には少なくとも一つ集積点が存在する . さらに , 任意の集積点はバリア KKT 条件を満足する .

したがって $\mu_k \rightarrow 0$ であることに注意すれば , 定理 2 よりアルゴリズム IP の大域的収束性が証明出来る .

4. 局所的超 1 次収束性

局所的収束性を調べるために , 次のアルゴリズムを考える (以下では簡単のために $p = n$ とする)

[アルゴリズム IPlocal]

Step 0. 初期点 $w_0 \in \mathbf{R}_+^n \times \mathbf{R}^m \times \mathbf{R}_+^n$ とパラメータ $\mu_0 > 0, 0 < \gamma_0 < 1, \varepsilon > 0$ を与える . $k = 0$ とおく .

Step 1. もし $\|r_0(w_k)\| \leq \varepsilon$ ならば停止する .

Step 2. パラメータ $\mu_k > 0, 0 < \gamma_k < 1$ を与える .

Step 3. ニュートン方程式

$$J(w_k) \Delta w_k = -r(w_k; \mu_k)$$

を解いて探索方向 Δw_k を求める .

Step 4. γ_k に対してステップ幅

$$\begin{aligned} \alpha_k = \min \left\{ 1, \gamma_k \min_i \left\{ -\frac{(x_k)_i}{(\Delta x_k)_i} \middle| (\Delta x_k)_i < 0 \right\}, \right. \\ \left. \gamma_k \min_i \left\{ -\frac{(z_k)_i}{(\Delta z_k)_i} \middle| (\Delta z_k)_i < 0 \right\} \right\} \end{aligned}$$

を計算する .

Step 5. $w_{k+1} = w_k + \alpha_k \Delta w_k$ とおく .

Step 6. $k := k + 1$ において ステップ 1 へ行く .

ここでは , 中心パスの近傍をたどりながら KKT 点に近づいて行った場合の点列の挙動について考える . 以下では , 反復回数 k は十分に大きいとし , パラメータ μ_k の値は十分に 0 に近いとし , さらに次の条件を仮定する .

仮定 CP

(CP1) 点列 $\{w_k\}$ は w^* に収束する .

(CP2) 関数 $f, g_i, i = 1, \dots, m$, のヘッセ行列が x^* においてリプシッツ連続である .

(CP3) w^* において1次独立制約想定，2次の十分条件，狭義相補条件が成り立つ．

(CP4) $\min(1, \tau_2) > \tau_1$, $0 < \kappa < 1$ となる正定数 τ_1, τ_2, κ に対して，パラメータ μ_k と γ_k が次式を満足する．

$$\mu_k = \xi_k \|r_0(w_k)\|^{1+\tau_1}, \quad 1 - \gamma_k = \kappa \xi_k \|r_0(w_k)\|^{\tau_2}$$

ただし， ξ_k は定数 M' に対して $\frac{1}{M'} \leq \xi_k \leq M'$ を満たす正数である．

(CP5) $0 < M_c < \sqrt{n}$.

(CP6) すべての k に対して，行列 G_k が

$$\|G_k - \nabla_x^2 L(w^*)\| < \delta,$$

$$\|(G_k - \nabla_x^2 L(w_k))\Delta x_k\| = O(\|\Delta w_k\|^{1+\tau_3})$$

を満足する．ただし， $\delta > 0$ は十分に小さな数であり， τ_3 は $\tau_3 > \tau_1$ を満たす正定数である．

以上の仮定のもとで次の定理を得る．

[定理 3]([14]) 仮定 CP が成り立つとする．このとき，アルゴリズム IPlocal で生成される点列 $\{w_k\}$ に対して次のことが成り立つ．

(1) もし点 $\hat{w} \in \mathbf{R}_+^n \times \mathbf{R}^m \times \mathbf{R}_+^n$ が $\|r(\hat{w}; \mu_k)\| \leq M_c \mu_k$ を満たすならば，適当な正定数 ν_1, ν_2 がとれて

$$\nu_1 \|r_0(w_k)\|^{1+\tau_1} \leq \|r_0(\hat{w})\| \leq \nu_2 \|r_0(w_k)\|^{1+\tau_1}$$

が成り立つ．

(2) もし $\|r(w_k; \mu_{k-1})\| \leq M_c \mu_{k-1}$ ならば，ステップ幅が $\alpha_k = 1$ となり，次式が成り立つ．

$$\|r(w_k + \Delta w_k; \mu_k)\| \leq M_c \mu_k$$

この定理によれば，もし w_k が μ_{k-1} に対して近似バリヤ KKT 条件を満たすならば，アルゴリズム IPlocal のステップ 4 で $\alpha_k = 1$ が採用されて，次の点 $w_k + \Delta w_k$ が新しいパラメータ μ_k に対する近似バリヤ KKT 条件を満足する．したがって，

$$\nu_1 \|r_0(w_k)\|^{1+\tau_1} \leq \|r_0(w_{k+1})\| \leq \nu_2 \|r_0(w_k)\|^{1+\tau_1}$$

が成り立つので，点列 $\{w_k\}$ の超 1 次収束性が保証される．

なお， (x, z) 空間における主双対内点法の局所的超 1 次収束性についての議論は Yabe and Yamashita[11] を参照されたい．

参考文献

- [1] I.Akrotirianakis and B.Rustem, *A globally convergent interior point algorithm for general nonlinear programming problems*, Technical Report 97-14, Department of Computing, Imperial College of Science, Technology and Medicine, March 1999.
- [2] M.Argaez and R.A.Tapia, *On the global convergence of a modified augmented Lagrangian linesearch interior point Newton method for nonlinear programming*, Technical Report TR95-38, Department of Computational and Applied Mathematics, Rice University, October 1995 (revised February 1997).

- [3] R.H.Byrd, J.C.Gilbert and J.Nocedal, *A trust region method based on interior point techniques for nonlinear programming*, Technical Report OTC 96/02, Optimization Technology Center, Argonne National Laboratory, June, 1996.
- [4] R.H.Byrd, M.E.Hribar and J.Nocedal, *An interior point algorithm for large scale nonlinear programming*, Technical Report OTC 97/05, Optimization Technology Center, Argonne National Laboratory, August, 1997.
- [5] R.H.Byrd, G.Liu and J.Nocedal, On the local behaviour of an interior point method for nonlinear programming, in *Numerical analysis 1997*, D.F.Griffiths, D.J.Higham and G.A. Watson eds., Longman (1998), pp.37-56.
- [6] J.E.Dennis, Jr., M.Heinkenschloss and L.N.Vicente, *Trust-region interior-point SQP algorithms for a class of nonlinear programming problems*, TR94-45, Dept. of Computational and Applied Mathematics, Rice University, Houston, Texas, USA, December 1994 (revised November 1995).
- [7] A.S.El-Bakry, R.A.Tapia, T.Tsuchiya and Y.Zhang, On the formulation and theory of the Newton interior-point method for nonlinear programming, *Journal of Optimization Theory and Applications*, 89 (1996), pp.507-541.
- [8] A.Forsgren and P.E.Gill, Primal-dual interior methods for nonconvex nonlinear programming, *SIAM J. on Optimization*, 8 (1998) pp.1132-1152.
- [9] H.J.Martinez, Z.Parada and R.A.Tapia, On the characterization of Q -superlinear convergence of quasi-Newton interior-point methods for nonlinear programming, *Bol. Soc. Mat. Mexicana*, 1 (1995), pp.137-148.
- [10] 矢部博, 八巻直一「非線形計画法」, 朝倉書店, 1999.
- [11] H.Yabe and H.Yamashita, Q -superlinear convergence of primal-dual interior point quasi-Newton methods for constrained optimization, *Journal of the Operations Research Society of Japan*, 40 (1997), pp.415-436.
- [12] H.Yamashita, A globally convergent primal-dual interior point method for constrained optimization, *Optimization Methods and Software*, 10 (1998), pp.443-469.
- [13] H.Yamashita and H.Yabe, Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization, *Mathematical Programming*, 75 (1996), pp.377-397.
- [14] H.Yamashita, H.Yabe and T.Tanabe, *A globally and superlinearly convergent primal-dual interior point trust region method for large scale constrained optimization*, Technical Report, Mathematical Systems, Inc., Tokyo, Japan, July 1997 (revised July 2000).
- [15] H.Yamashita and H.Yabe, *An interior point method with a primal-dual l_2 barrier penalty function for nonlinear optimization* Technical Report, Mathematical Systems, Inc., Tokyo, Japan, March 1999.