

非線形計画法 (4)

—制約条件付き最適化問題—

矢部 博, 八巻 直一

1. はじめに

今回は、制約条件のついた最適化問題の数値解法を紹介します。この問題は、次の形で定式化されます。

制約条件付き問題

n 変数の非線形関数 $f(x)$ (目的関数) を、等号制約条件

$$h_i(x) = 0, \quad i = 1, \dots, m$$

と不等号制約条件

$$g_j(x) \leq 0, \quad j = 1, \dots, l$$

のもとで、最小化せよ。

ここで、 $f(x)$ を目的関数、 $g_j(x)$ や $h_i(x)$ を制約関数とよびます。とくに、関数 $f(x), g_j(x), h_i(x)$ のすべてが x についての 1 次関数のとき、線形計画(LP)問題と呼び、 $f(x)$ が 2 次関数で $g_j(x), h_i(x)$ が 1 次関数のとき 2 次計画(QP)問題と呼びます。さらにいずれかの関数が非線形であるとき、総称して非線形計画問題と呼びます。

LP 問題は、それ専用の有力な数値解法がほぼ確立されていて、第 2 次世界大戦が終結して間もない 1947 年にアメリカの G.B.Dantzig が発表したシンプレックス法(単体法)が、47 年を経た今日でも不動の地位を築いていることは周知の事実です。しかしながらその一方で、LP 問題の数値解法の見直しもなさ

れ、アメリカの AT & T ベル研究所の N.Karmarkar が 1984 年に発表した内点法が近年注目を浴びています。これはとりわけ、「数学的手法が特許申請された例」としても有名です。LP 問題での内点法の成功がきっかけとなって、現在、非線形計画問題における内点法の見直しがなされつつあります。

以下で述べる解法は反復法です。これは、与えられた初期点から出発して解の近似点を生成していくものです。その際に、初期点をいかにうまく選ぶかが重要な問題になります。解の十分近くに初期点を選べば収束性が保証される場合、その数値解法は局所的に収束するといいい、この場合には収束速度が議論されます。一方、解から離れた初期点から出発しても解に収束する場合、その解法は大域的収束するといいい、この場合は数値解法の頑健性が議論されます。現在、大域的収束を実現する手段として、探索方向での刻み幅を調整する直線探索法と、解の近似点のまわりの限られた領域で次の点を探す信頼領域法の 2 つが代表的です。

具体的な数値解法を紹介する前に、まず準備として最適性条件を述べましょう。以下では、ベクトル $v = (v_1, \dots, v_n)^T$ に対して $\|v\|$ は l_2 ノルムを表し、

$$\|v\| = \sqrt{\sum_{i=1}^n v_i^2}$$

で与えられます。

また、

$$g(x) = (g_1(x), \dots, g_l(x))^T,$$

$$h(x) = (h_1(x), \dots, h_m(x))^T$$

と定義します。

制約条件を満たす点の集合

$$S = \{x \in R^n \mid g(x) \leq 0, h(x) = 0\}$$

やべ ひろし 東京理科大学工学部

〒162 新宿区神楽坂 1-3

やまき なおかず システム計画研究所

〒150 渋谷区桜丘町 2-9 カサヤビル

を許容領域といい、 S に属する点を許容解といいます。任意の許容解 x に対して、 $f(x_*) \leq f(x)$ が成り立つとき、許容解 $x_* \in S$ は大域的最適解であるといいます。また、 x_* の近傍内の任意のベクトル x に対して $f(x_*) \leq f(x)$ が成り立つとき、許容解 $x_* \in S$ は局所的最適解であるといいます。

ラグランジュ関数を

$$\begin{aligned} L(x, \lambda, \mu) &= f(x) + \lambda^T g(x) + \mu^T h(x) \\ &= f(x) + \sum_{j=1}^l \lambda_j g_j(x) + \sum_{i=1}^m \mu_i h_i(x) \end{aligned}$$

と定義したとき、制約条件付き問題を解くことは次の5つの条件を満足する点 (x_*, λ_*, μ_*) を見つけることに帰着されます。

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= \nabla f(x) + \sum_{j=1}^l \lambda_j \nabla g_j(x) \\ &\quad + \sum_{i=1}^m \mu_i \nabla h_i(x) = 0 \\ \nabla_\lambda L(x, \lambda, \mu) &= g(x) \leq 0, \\ \nabla_\mu L(x, \lambda, \mu) &= h(x) = 0, \\ \lambda_i &\geq 0, \quad i = 1, \dots, l, \\ \lambda_i g_i(x) &= 0, \quad i = 1, \dots, l. \end{aligned}$$

ただし、 λ, μ をそれぞれ不等号、等号条件に対するラグランジュ乗数（ベクトル）といい、また、最後の条件式を相補条件といいます。

以上の5つの条件をカルーシュ・キューン・タッカー（KKT: Karush-Kuhn-Tucker）条件といいます。とくに、 $f, g_j, j = 1, \dots, l$ がすべて凸関数で、 $h_i, i = 1, \dots, m$ のすべてが線形ならば KKT 点の x_* は最適解になることが証明されています。

こうした裏付けのもとで、既存の数値解法の多くは KKT 条件を満足する点を求めることを目指しているわけです。

2. ペナルティ法

それでは、代表的な数値解法を紹介していきましょう。まず最初に、本節ではペナルティ法をいくつか紹介します。

2.1 内点ペナルティ法

不等号条件だけが付いた問題を考えましょう。内点ペナルティ法 (interior penalty function method または barrier function method) の原理は、許容領域内での関

数値が、領域の境界に近づくにつれて大きくなり、境界上では無限大になるような関数を定義し、その関数の最小点を無制約最小化法を用いて求めることにあります。このような関数をバリアー関数といいます。不等号条件に対するバリアー関数として、次のものが考えられています。

$$\begin{aligned} P(x; \tau) &= f(x) - \tau \sum_{i=1}^l \frac{1}{g_i(x)}, \\ P(x; \tau) &= f(x) - \tau \sum_{i=1}^l \log(-g_i(x)). \end{aligned}$$

とくに後者はログバリアー関数と呼ばれるもので、LP の新解法では重要な役割を果たしています。

内点ペナルティ法のアルゴリズムは、まず初期点 x_0 （許容領域の内点）を選んで、単調に減少して $\tau_k \rightarrow 0$ となるペナルティパラメータ列 $\{\tau_k\}$ に対して、逐次バリアー関数 $P(x; \tau_k)$ を最小にする点 x_k を求めて、最適解の近似点列 $\{x_k\}$ を生成していくものです。その際、バリアー関数の最小化には、前回紹介した無制約最小化法が利用されます。

このとき、適当な仮定のもとで

$$\lim_{k \rightarrow \infty} P(x_k; \tau_k) = f(x^*), \quad \lim_{k \rightarrow \infty} f(x_k) = f(x^*)$$

が成り立ちます。

実際の問題では、最適解を得ることもさることながら、より良い許容解を得ることが当面の目的である場合が多いので、内点ペナルティ法はいつ停止しても初期点よりも良い許容解が得られている、という利点を持っています。しかしながら、制約条件をみだす初期点を必要とすることや、許容領域の境界に近づくにつれて数値的に不安定になることなどが、問題点として挙げられます。

2.2 外点ペナルティ法と混合法

最小化問題に等号条件が含まれる場合には、もはや内点ペナルティ法は使えません。このことを克服するために外点ペナルティ法 (exterior penalty method) が考えられています。この方法は、等号条件を考慮し得るほか、初期点として必ずしも内点を求めなくてもよいように考案されています。

外点ペナルティ関数は x の全域で定義され、許容領域内ではペナルティ項が0、許容領域の外側では境界から離れるに従って、関数値が無限大に近づくように構成されます。

外点ペナルティ関数の例として

$$\nabla_x L(x, \mu) = 0, \nabla_\mu L(x, \mu) = 0 \quad (2)$$

となります。このとき、この方程式にニュートン法を適用すれば k 回目における反復式は

$$x_{k+1} = x_k + \Delta x_k, \mu_{k+1} = \mu_k + \Delta \mu_k$$

で与えられます。ただし、 Δx_k と $\Delta \mu_k$ は補正ベクトルで、次の連立1次方程式の解です。

$$\nabla^2 L(w_k) \begin{pmatrix} \Delta x_k \\ \Delta \mu_k \end{pmatrix} = - \begin{pmatrix} \nabla_x L(w_k) \\ h(x_k) \end{pmatrix}, \quad (3)$$

$$\nabla^2 L(w_k) = \begin{pmatrix} \nabla_{xx} L(w_k) & \nabla h(x_k) \\ \nabla h(x_k)^T & 0 \end{pmatrix}.$$

ここで $\nabla h(x) = [\nabla h_1(x), \dots, \nabla h_m(x)] \in R^{n \times m}$ は $h(x)$ のヤコビ行列の転置行列、 $\nabla_{xx} L(x, \mu)$ は $L(x, \mu)$ の x に関するヘッセ行列、および $w_k = (x_k, \mu_k)$ です。

乗数法の原理は、上記のようにラグランジュ関数を構成して、その停留点を探索することにあります。もし、ラグランジュ関数のヘッセ行列 $\nabla^2 L$ が正定値行列であれば、ニュートン法は安定に極値に収束することが期待できます。しかし、残念なことに $\nabla^2 L$ は正定値行列ではないのでそのままニュートン法を適用することには問題があります。そこで、ラグランジュ関数にペナルティ項を付加して、局所的に関数が凸になるように工夫することが、Hestenes(1969) と Powell(1969) によって独立に試みられました。この関数を拡張ラグランジュ関数 (augmented Lagrangian function) と呼び、次のように与えられます。

$$Q(x, \mu; r) = f(x) + \sum_{i=1}^m \mu_i h_i(x) + \frac{1}{2} r \sum_{i=1}^m h_i(x)^2$$

ただし、 r はペナルティパラメータです。もし $\nabla Q = 0$ となる点が制約条件をみたすならば、それはラグランジュ関数の停留点になります。乗数法の重要な利点のひとつは、ペナルティパラメータ r が有限の値でよいということです。

乗数法のアルゴリズムの基本は、与えられた μ_k, r_k に対して、無制約最小化法で関数 $Q(x, \mu_k, r_k)$ の最小点 x_k を求めて点列 $\{x_k\}$ を生成することです。その際、ペナルティパラメータ r_k とラグランジュ乗数の推定値 μ_k が更新されます。ラグランジュ乗数の推定値の更新については、いくつかの更新公式が提案されていますが、次の2つが代表的です。

$$\begin{aligned} \mu_{k+1} &= \mu_k + r_k h(x_k), \\ \mu_{k+1} &= \mu_k + H_k^{-1} h(x_k) \end{aligned}$$

$$P(x; r) = f(x) + r \left\{ \sum_{j=1}^l |\max(g_j(x), 0)|^\alpha + \sum_{i=1}^m |h_i(x)|^\beta \right\}$$

があります。ここに、 $\alpha, \beta \geq 1$ 、 r は正のペナルティパラメータで、 $r \rightarrow \infty$ の場合を考えます。

さらに、不等号条件に対してはバリアー関数の性質をもち、等号条件に対しては外点ペナルティ関数の性質をもつペナルティ関数を作ることができます。これを混合ペナルティ関数 (mixed penalty function) といひ、たとえば、

$$P(x; r) = f(x) - r \sum_{j=1}^l \frac{1}{g_j(x)} + \frac{1}{r^2} \sum_{i=1}^m h_i(x)^2$$

が考えられます。

なお、Fiacco and McCormick(1968) のSUMT法は、一般の制約付き問題に対する混合ペナルティ法のことです。

2.3 正確なペナルティ関数法

これまで述べたペナルティ法では、ペナルティパラメータ r を $r \rightarrow 0$ (もしくは $r \rightarrow \infty$) としながら無制約最小化 (部分) 問題を何回も解かなければなりません。これに対して、ペナルティパラメータを変更することなく、変換された無制約最小化問題を1回解くだけで、もとの問題の解が得られるような手法が考えられています。これは正確なペナルティ関数法 (exact penalty function method) と呼ばれ、Zangwill(1967) によって提案されました。正確なペナルティ関数として、たとえば

$$P(x; r) = f(x) + r \left\{ \sum_{j=1}^l \max(g_j(x), 0) + \sum_{i=1}^m |h_i(x)| \right\} \quad (1)$$

がよく知られています。

3. 乗数法

ペナルティ関数を用いる方法は、許容領域の境界付近で数値的に不安定になる、という問題点を抱えています。ペナルティ法の持つこの欠点を回避する方法として、乗数法 (multiplier method) があげられます。等号条件付き最小化問題を考えてみましょう。ラグランジュ関数を $L(x, \mu)$ としたとき、KKT条件は

ここに,

$$H_k = \nabla h(x_k)^T \nabla_{xx} Q(x_k, \mu_k; \tau_k)^{-1} \nabla h(x_k)$$

です.

不等号条件がある場合には, スラック変数を導入して, 以上と同様の議論ができます.

4. 逐次2次計画法

本章では, 準ニュートン法の考え方に基づく方法を紹介します. まず等号条件付き問題を考えてみましょう. 基本的な考え方は乗数法と同じで, 非線形方程式(2)に対するニュートン法です. ただし, もうひと工夫します. すなわち, $\mu_{k+1} = \mu_k + \Delta\mu_k$ とおくと, 方程式(3)は

$$\nabla^2 L(x_k, \mu_k) \begin{pmatrix} \Delta x_k \\ \mu_{k+1} \end{pmatrix} = - \begin{pmatrix} \nabla f(x_k) \\ h(x_k) \end{pmatrix}$$

と書けます.

ここでさらに, 無制約問題における準ニュートン法の考え方を用いて, ヘッセ行列 $\nabla_{xx} L(x_k, \mu_k)$ を適当な近似行列 B_k で置き換えれば, 上式は次の方程式に帰着されます.

$$B_k \Delta x_k + \nabla f(x_k) + \nabla h(x_k) \mu_{k+1} = 0,$$

$$h(x_k) + \nabla h(x_k)^T \Delta x_k = 0.$$

実は, この式が次のQP問題のKKT条件であることが容易に確かめられます.

QP問題(a)

線形等号条件

$$h(x_k) + \nabla h(x_k)^T \Delta x = 0$$

のもとで, 2次関数

$$\frac{1}{2} \Delta x^T B_k \Delta x + \nabla f(x_k)^T \Delta x$$

を $\Delta x \in R^n$ について最小にせよ.

このとき B_k が正定値対称行列ならば, このQP問題は一意解 Δx_k をもち, μ_{k+1} が線形等号条件に対するラグランジュ乗数ベクトルになります. したがって, 非線形方程式を解く代わりにQP問題(a)を解いて探索方向 Δx_k を求める解法が考えられるわけです.

以上の考え方をそのまま制約条件付き問題に適用すれば, 次の問題が得られます.

QP問題(b)

線形制約条件

$$g(x_k) + \nabla g(x_k)^T \Delta x \leq 0,$$

$$h(x_k) + \nabla h(x_k)^T \Delta x = 0$$

のもとで, 2次関数

$$\frac{1}{2} \Delta x^T B_k \Delta x + \nabla f(x_k)^T \Delta x$$

を $\Delta x \in R^n$ について最小にせよ.

以上のように, 各反復でQP問題を解いて探索方向 Δx_k を決定していく方法を逐次2次計画法 (sequential quadratic programming method) といいます. この方法は, ヘッセ行列 $\nabla_{xx} L(x_k, \lambda_k, \mu_k)$ を行列 B_k で近似するという意味から, 制約条件付き問題に対する準ニュートン法とみなすことができます.

準ニュートン法という観点からみれば, B_k の更新式をどう選ぶかが重要な問題になります. これに関しては, Powell(1978)が正定値性を保存することを考慮して, BFGS公式に対応する次の更新式を提案しました.

$$B_{k+1} = B_k - \frac{B_k s_k (B_k s_k)^T}{(B_k s_k)^T s_k} + \frac{z_k z_k^T}{z_k^T s_k}$$

ただし,

$$z_k = \phi_k y_k + (1 - \phi_k) B_k s_k,$$

$$s_k = x_{k+1} - x_k,$$

$$y_k = \nabla_x L(x_{k+1}, \lambda_{k+1}, \mu_{k+1}) - \nabla_x L(x_k, \lambda_{k+1}, \mu_{k+1})$$

$$\phi_k = \begin{cases} y_k^T s_k \geq \psi s_k^T B_k s_k \text{ のとき} & 1 \\ \text{そうでないとき} & \frac{(1-\psi) s_k^T B_k s_k}{s_k^T (B_k s_k - y_k)} \end{cases}$$

です. ここで, $\psi = 0.1$ または $\psi = 0.2$ が選ばれます. とくに $\phi_k = 1$ が採用された場合には, 本来のセカント条件を満足するBFGS公式になります. この更新式をPowellの修正BFGS公式といいます.

収束性に関しては, 正確なペナルティ関数(1)に直線探索を適用した逐次2次計画法の大域的収束性がHan(1977)によって示されています. また, 収束の速さに関して, Powell(1978)が修正BFGS公式の超1次収束性を論じています.

5. 主双対内点法

Karmarkar 法以来, LP 問題や QP 問題に対する内点法が理論的にも実用的にも非常に活発に研究されています. LP 問題に対する内点法の中でも近年とくに主双対内点法 (primal-dual interior point method) が有望視されており, Kojima, Mizuno and Yoshise(1989) の研究に端を発して, 多くの研究者によって大域的収束性(とくに多項式オーダー性)や局所的収束性が研究されています. そうした状況の中で, 非線形最適化問題の内点法の検討は今後の大きな課題です. 本節では, 非線形最適化問題に対する主双対内点法を紹介します.

非線形最適化問題は次の形に書き換えられることに注意しましょう.

$$h(x) = 0, x \geq 0 \text{ のもとで } f(x) \text{ を最小化せよ}$$

不等号条件 $g(x) \leq 0$ はスラック変数を導入すれば, 上式の形で書けます. 上記の問題のラグランジュ関数を

$$L(w) = f(x) - y^T h(x) - z^T x \quad (4)$$

としたとき, KKT 条件は次式で与えられます:

$$\nabla_x L(w) = 0, h(x) = 0, XZe = 0,$$

$$x \geq 0, z \geq 0, \text{ (非負条件)}$$

ただし, $y \in R^m, z \in R^n$ はそれぞれ等号条件, 非負条件に対するラグランジュ乗数, $w = (x, y, z)^T, X = \text{diag}(x_1, x_2, \dots, x_n), Z = \text{diag}(z_1, z_2, \dots, z_n), e = (1, 1, \dots, 1)^T \in R^n$ です. このとき相補条件 $XZe = 0$ を $XZe = \tau e$ (τ は非負定数) で置き換えた次の方程式を考えます:

$$r(w) \equiv \begin{pmatrix} \nabla_x L(w) \\ h(x) \\ XZe - \tau e \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad (5)$$

$$x > 0, z > 0. \text{ (} x, z \text{ が正であることに注意)}$$

主双対内点法は, この非線形方程式に対してニュートン法を適用したものです. 主双対内点法では, 変数 (x_k, z_k) が非負条件に関して内点になるように刻み幅が調整されます. すなわち,

$$x_{k+1} = x_k + \alpha_{xk} \Delta x_k > 0,$$

$$z_{k+1} = z_k + \alpha_{zk} \Delta z_k > 0$$

となるように α_{xk}, α_{zk} が選ばれます. したがって, 主双対内点法と従来のニュートン法との本質的な違いは, 扱う方程式に摂動項 $\tau_k e$ が含まれていること, 点列 $\{(x_k, z_k)\}$ が非負条件に関して内点になるように減速をしなければならないこと, です. さらに, このことに付け加えて, ラグランジュ関数のヘッセ行列 $\nabla_{xx} L(w_k)$ を近似することも考えれば, ニュートン方程式に対応する連立1次方程式 $J_k \Delta w_k = -r(w_k)$ を解いて探索方向 $\Delta w_k = (\Delta x_k, \Delta y_k, \Delta z_k)^T$ を求める解法が得られます. ただし,

$$J_k = \begin{pmatrix} G_k & -A(x_k)^T & -I \\ A(x_k) & O & O \\ Z_k & O & X_k \end{pmatrix}$$

です. ここで $G_k = \nabla_{xx} L(w_k)$ の場合がニュートン法に基づく主双対内点法, G_k をヘッセ行列の近似行列にとった場合が準ニュートン法に基づく主双対内点法になります. 収束性に関しては, 非負条件にログバリアー関数を, そして等号条件に正確なペナルティ関数をそれぞれ適用して得られる直線探索評価関数と Armijo の直線探索法を組み合わせることによって, Yamashita(1992) が大域的収束性を示しています. また, 局所的超1次収束性が Yamashita and Yabe(1993) によって示されています.

6. おわりに

以上で, 制約付き問題に対する代表的な数値解法の紹介を終わります. ここで述べた解法以外にも, 信頼領域法や射影法など有効な数値解法がいくつかありますし, また, コンプレックス法のような微分の情報を使わない解法もありますが, ここでは割愛しました. 実際問題としてよく扱われる, 制約条件付き非線形最小二乗問題に対する数値解法にも触れませんでした.

さて, 非線形計画法に関する連載も今回で終了いたします. 本連載の主なねらいは, 第一回目で取り上げた専門書の紹介です. 残りの三回分はそのための補足であり, これから非線形計画法を勉強しようとする人たちが気軽に専門書をひもとくためのウォーミングアップを兼ねています. したがって, ここで十分に説明できなかった方法については, 第一回で紹介した専門書を参考にしていただければ幸いです.