数理解析研究所講究録
680巻 1989年 9-18

非線形計画問題の分離的凸化による逐次解法

金沢大・工　谷川　明夫（Akio Tanikawa）

1. 序

次のような等式制約条件をもつ非線形計画問題について考える：

$$(1.1)\quad \begin{cases} \text{目的関数} & f(x) \longrightarrow \text{最小} \\ \text{制約条件} & g(x) = 0, \quad x \in D \end{cases}$$

ただし，$f : \mathbb{R}^n \longrightarrow \mathbb{R}$，$g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$ で D は $\mathbb{R}^n$ の有界な開集合とする．すなわち，局所的な等式制約条件付非線形計画問題を考える．とくに応用上重要な大規模計画問題（次元 n が大きい場合）の多くは次の形の変数に関する分離性を有する：

$$(1.2)\quad \begin{cases} f(x) = \sum_{k=1}^{p} f_k(\xi^k) \\ g^j(x) = \sum_{k=1}^{p} g_k^j(\xi^k), \quad j = 1, 2, \cdots\cdots, m. \end{cases}$$

ただし，$g(x) = (g^1(x), g^2(x), \cdots\cdots, g^m(x))$，$x = (\xi^1, \xi^2, \cdots, \xi^p)$，

$\xi^k \in \mathbb{R}^{n_k}$, $n_1 + n_2 + \cdots + n_p = n$. そこでここでは，分離性を有する高次元の最小化問題を低次元の最小化問題に分解して解く効率的な手順を考察する.

2. 主－双対法とラグランジュ関数の凸化

分離的な大規模計画問題を解く手順としては，主－双対法が有力な方法としてよく知られている．主－双対法は問題(1.1) が局所凸条件をみたせば，その最適解 x^* とそれに対応するラグランジュ乗数 y^* がラグランジュ関数 $l: \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}$

$$(2.1) \qquad l(x, y) = f(x) + y^T g(x)$$

の局所的な鞍点になるという性質に基づき，（制約条件付）最小化問題（1.1）を一連の（無制約条件）最小化問題の解法により求める方法である（[6] 参照）．この方法の長所は，双対関数 $g(y)$

$$(2.2) \qquad g(y) = \min_x \ l(x, y)$$

を求めるに際し，(1.2) の条件の下では，低次元の最小化への分解

$$(2.3) \qquad \min l(x, y) = \sum_{k=1}^{p} \min_{\xi^k} \left\{ f_k(\xi^k) + y^T g_k(\xi^k) \right\}$$

が成り立つことである．しかし一般には $l(x, y)$ は x について局所凸にならない（i.e., 問題 (1.1) が局所凸条件をみたさ

ない）ので，この方法の適用範囲はあまり広くない．

そこで，$l(x, y)$ を局所凸化する方法が提案されたが，最も有名なものは乗数法である．この方法は (1.1) と同値な問題

(2.4) $$\min \left\{ f(x) + (C/2)\|g(x)\|^2 \,\middle|\, g(x) = 0,\ x \in D \right\}$$

を考えれば，これに対応するラグランジュ関数 $l(x, y)$ は，C が十分大きいとき，x について局所凸になるという性質に着目している（[3], [4], [7] 参照）．しかしこの方法の欠点は，f, g が分離性の条件 (1.2) をみたすときでも，(2.4) に対応するラグランジュ関数 $l(x, y)$ はもはや x について分離的ではなく，(2.3) のような低次元への分解が成り立たないことである．

その後，上の欠点を補ういくつかの方法（[1]，[8], [12] 参照）が提案されたが，ここでは谷川・向井 [9], [10]，[11] の考えに沿った逐次解法について考える． [10], [11] では，次の形の分離的に凸化されたラグランジュ関数

(2.5) $$l_{c,b}(x, z) = l(x, y(z)) + b\, g(z)^T M(z) g(x) + (C/2)\|z - x\|^2$$

を考察した．ただし，$b \geqq 0$，$C \geqq 0$，$M(z)$ は正値対称行列，そして $y(z)$ は Fletcher [2] で定義された関数

(2.6) $$y(x) = \arg\min_{y} \|\nabla_x l(x, y)\|$$

である．ここで $\nabla_x l(x, y)$ は $l(x, y)$ の x についての勾配ベク

トル（縦ベクトル）を表わす．$l_{c,b}(x,z)$ は x について (2.3) と同様な性質──低次元への分解──を有する．$l_{c,b}(x,z)$ は [9] では次の形に一般化された：

$$(2.7)\qquad L_{c,b}(x,z) = \alpha\, l(x,y(z)) + \beta\, \nabla_x l(z,y(z))^T \nabla_x l(x,y(z)) + b\,g(z)^T M(z)\, g(x) + (c/2)\|x-z\|^2.$$

そして，逐次過程 $\{z_k\}$ を

$$(2.8)\qquad z_{k+1} = \arg\min_x L_{c,b}(x, z_k)$$

で定義すれば，$\alpha+\beta>0$ という条件の下で x^* に局所収束（１次平均収束）することが示された．

ここでは，その特別な場合「$\alpha=0$，$\beta=1$」について $L_{c,b}(x,z)$ の性質と逐次解法について考える．また結果を簡略化するために，b，$M(z)$ は次の形のものを考える：

$$(2.9)\qquad b = c \quad ; \quad M(z) = [(\partial g(z)/\partial x)(\partial g(z)/\partial x)^T]^{-1}.$$

ここで，$(\partial g(x)/\partial x)$ は横ベクトル $\nabla g^j(x)^T$，$j=1,2,\cdots,m$ から成る $m\times n$ ヤコビ行列とする．このとき [11] と同様な結果が得られることを以下で示す．すなわち関数 $F_c(z)$

$$(2.10)\qquad F_c(z) = \min_x L_{c,c}(x,z)$$

は，c が十分大きければ，原問題 (1.1) の最適解 x^* で（局所的な）最小値をとる局所凸関数になる．

3．定義と結果

まず以下の仮定をおく．

仮定(I)：　$f : D(\subset \mathbb{R}^n) \longrightarrow \mathbb{R}$，　$g : D \longrightarrow \mathbb{R}^m$　は C^3 級関数である．

仮定(II)：　$(x^*, y^*) \in \mathbb{R}^n \times \mathbb{R}^m$ は局所最適解のための２次の十分条件をみたす：

$$(3.1)\quad \begin{cases} \nabla_x \ell(x^*, y^*) = 0, & g(x^*) = 0 \\ u^T \nabla_x^2 \ell(x^*, y^*)\, u > 0 & ({}^\forall u : (\partial g(x^*)/\partial x)u = 0). \end{cases}$$

仮定(III)：　$\nabla g^j(x)$，$j = 1, 2, \cdots, m$　は任意の $x\ (\in D)$ に対して１次独立である．

仮定(IV)：　$L_{c,b}(x, z)$　において $\alpha = 0$，$\beta = 1$ であり，b，$M(z)$　は（2.9）で与えられる．

注意：　仮定(III)によって，(2.6)で定義された $y(x)$ は任意の $x\ (\in D)$ に対して唯１つ存在する（[10]）．

十分大きな c に対して，$L_{c,c}(x, z)$ が x について局所凸であることは $\ell_{c,b}(x; z)$ と同様である．したがって，

$$(3.2)\qquad \Phi(z) = \arg\min_x\ L_{c,c}(x, z)$$

も定まる．$\Phi(z)$ は記号の簡略化のために $\hat{z}$ とも書く．明らかに $\hat{x}^* = x^*$ が成り立つ．また以下では記号 $\nabla g(z) = (\partial g(z)/\partial x)^T$ の方を主に用いる．次の命題は[11]の Proposition 3 と同様に初

等的な計算により得られる．

<u>命題 3.1</u>： 次の等式が成り立つ：

$$\nabla F_c(z) = -c(\hat{z}-z) + c\,\nabla g(z) M(z) g(\hat{z}) + \nabla y(z) \nabla g(\hat{z})^T \nabla_x l(z, y(z)) + \left[\nabla_x^2 l(z, y(z)) + \nabla y(z) \nabla g(z)^T\right] \nabla_x l(\hat{z}, y(z)) + c \sum_{i,j} g^i(z) g^j(\hat{z}) \nabla M_{ij}(z). \tag{3.3}$$

したがってとくに

$$\nabla F_c(x^*) = 0 \tag{3.4}$$

が成り立つ．ただし，∇^2は Hessian，$M_{ij}(z)$ は行列 $M(z)$ の (i,j) 成分を表わす．

ここで制約条件に対する点x^*における接部分空間を$T(x^*)$と表わせば，次のように書ける：

$$T(x^*) = \{u \in \mathbb{R}^n \mid \nabla g(x^*)^T u = 0\}. \tag{3.5}$$

このとき，接部分空間$T(x^*)$への射影行列は次の形になる：

$$P_1 = I - P_2. \tag{3.6}$$

ただし，P_2 は $T(x^*)$ の直交補空間 $T(x^*)^{\perp}$ への射影行列で

$$P_2 = \nabla g(x^*) \left[\nabla g(x^*)^T \nabla g(x^*)\right]^{-1} \nabla g(x^*)^T \tag{3.7}$$

が成り立つ．また以下では記号の簡略化のため

$$A = \nabla_x^2 l(x^*, y^*) \tag{3.8}$$

とする．

次の等式は [11] の Proposition 4 と同様に証明できる．

<u>命題 3.2</u>： 次の等式が成り立つ：

$$(3.9)\qquad \nabla\Phi(x^*) = P_1 - (1/c)\,AP_1A.$$

命題 3.1，命題 3.2 より次が得られる．

<u>命題 3.3</u>： 次の等式が成り立つ：

$$(3.10)\qquad \nabla^2 F_c(x^*) = cP_2 + AP_1AP_1 + P_1AP_1A - c^{-1}AP_1A^2P_1A.$$

命題 3.3 より，初等的な計算により次の定理を証明できる．

<u>定理</u>： ある正数 c_0 が存在して，$c \geqq c_0$ ならば $\nabla^2 F_c(x^*)$ は正定値である．

また c_0 についての上からの評価として次のものは比較的容易に得られる：

$$(3.11)\qquad c_0 \leqq (\|A\|^2/\lambda)^2 + \|A\|^2.$$

ただし，$\lambda > 0$ は次の不等式

$$(3.12)\qquad \langle AP_1u, P_1u\rangle \geqq \lambda\|P_1u\|^2, \qquad \forall u \in R^n$$

をみたすもので，これは仮定(II)より保障されている．ここで $\langle\ ,\ \rangle$ は R^n 上の内積を表わす．

命題 3.1 の (3.4) と定理より，関数 F_c は（局所）最適解 x^* の近傍で凸であることがわかる．したがって，F_c を最小

にするためには最急降下法

$$(3.13)\qquad z_{k+1} = z_k - \alpha_k \nabla F_c(z_k)$$

またはニュートン法

$$(3.14)\qquad z_{k+1} = z_k - \left[\nabla^2 F_c(z_k)\right]^{-1} \nabla F_c(z_k)$$

などを用いればよい．ここで，z_k が x^* に近いときは $\|g(z_k)\|$，$\|g(\hat{z}_k)\|$ が非常に小さくなり (3.3) より近似的に

$$(3.15)\qquad \nabla F_c(z_k) \simeq -c\,(\hat{z}_k - z_k)$$

となるので，$\alpha_k = c^{-1}$ ととれば (3.13) は [9] で定義された

$$(3.16)\qquad z_{k+1} = \hat{z}_k$$

という逐次解法になる．また実際の問題に対しては，例えば次のような折衷的な逐次解法も考えられる

$$(3.17)\qquad z_{k+1} = \begin{cases} \hat{z}_k & \|z_k - z_{k-1}\| \geqq 1 \\ (3.14) & \|z_k - z_{k-1}\| < 1 \end{cases}$$

ただし，(3.15) における $\nabla^2 F_c(z_k)$ は (3.10) 式の右辺の z_k における値を用いる．

次に簡単な例について逐次解法 (3.16) と (3.17) をそれぞれ用いたときの数値結果を以下に示す．

<u>例題</u>（[5] の Problem 5 (P.397) の非負条件を削除したもの）

$$f(x) = 1000 - (x^1)^2 - 2(x^2)^2 - (x^3)^2 - x^1x^2 - x^1x^3,$$

$$g^1(x) = (x^1)^2 + (x^2)^2 + (x^3)^2 - 25,$$

$$g^2(x) = 8x^1 + 14x^2 + 7x^3 - 56,$$

ただし，$x=(x^1,x^2,x^3)$ とする．次の表は $z_0=(13,6,13)$ から出発して逐次解法が $\|z_k-z_{k-1}\|<10^{-3}$ をみたすのに要した反復の回数 k を示したもので，Table 1は (3.16)，Table 2は (3.17) に対する結果である．また，(3.10)式の右辺の最後の項は削除してもほぼ同じ結果が得られることもわかった．参考までに，[5] によれば，例題の解は次の通りである：

$$f(x^*)=961.715,\quad x^{*1}=3.512,\quad x^{*2}=0.217,\quad x^{*3}=3.552.$$

Table 1

c	Iter. (k)	$f(z_k)$	$g^1(z_k)$	$g^2(z_k)$	z^1_k	z^2_k	z^3_k
5	16	961.715	2×10^{-5}	-3×10^{-3}	3.5116	0.2168	3.5528
10	23	961.716	-2×10^{-4}	-4×10^{-3}	3.5098	0.2169	3.5541
50	44	961.716	-1×10^{-4}	-5×10^{-5}	3.4978	0.2178	3.5662
100	42	961.717	-9×10^{-5}	-5×10^{-3}	3.4824	0.2191	3.5812
200	9	961.721	-2×10^{-4}	-4×10^{-3}	3.4569	0.2215	3.6056

Table 2

c	Iter. (k)	$f(z_k)$	$g^1(z_k)$	$g^2(z_k)$	z^1_k	z^2_k	z^3_k
5	12	961.715	0	-1×10^{-3}	3.5119	0.2169	3.5524
10	13	961.715	-5×10^{-5}	-6×10^{-4}	3.5118	0.2170	3.5524
50	12	961.715	0	-2×10^{-4}	3.5117	0.2170	3.5526
100	12	961.715	-2×10^{-5}	-9×10^{-5}	3.5117	0.2170	3.5526
200	12	961.715	0	-5×10^{-5}	3.5117	0.2170	3.5526

References

[1] D.P.Bertsekas, "Convexification procedures and decomposition methods for nonconvex optimization problems," J. Optimiz. Theory Appl., Vol. 29 (1979) 169-197.

[2] R.Fletcher, "A class of methods for nonlinear programming with termination and convergence properties," in Integer and Nonlinear Programming, J.Abadie, Ed., Amsterdam, The Netherlands: North Holland (1970) 157-175.

[3] P.C.Haarhoff and J.Buys, "A new method for the optimization of a nonlinear function subject to nonlinear constraints," J. Comput., Vol. 13 (1970) 178-184.

[4] M.R.Hestenes, "Multiplier and gradient methods," J. Optimiz. Theory Appl., Vol. 4 (1969) 303-320.

[5] D.M.Himmelblau, Applied Nonlinear Programming, New York New York: McGraw-Hill (1972).

[6] D.G.Luenberger, Introduction to Linear and Nonlinear Programming, Reading, MA: Addison-Wesley (1973).

[7] M.J.D.Powell, "A method for nonlinear constraints in minimization problems," in Optimization, R.Fletcher, Ed., New York, New York: Academic (1969).

[8] G.Stephanopoulos and W.Westerberg, "The use of Hestenes' method of multipliers to resolve dual gaps in engineering system optimization," J. Optimiz. Theory Appl., Vol. 15 (1975) 285-309.

[9] A.Tanikawa, "On a generalized class of separable convexification procedures in nonlinear programming," J. Toyota College of Technology, Vol. 20 (1987) 125-132.

[10] A.Tanikawa and H.Mukai, "A new technique for nonconvex primal-dual decomposition of a large-scale separable optimization problem," IEEE Trans. Autom. Control, Vol. AC-30 (1985) 133-143.

[11] A.Tanikawa and H.Mukai, "New Lagrangian function for nonconvex primal-dual decomposition," Comput. Math. Appl., Vol. 13 (1987) 661-676.

[12] N.Watanabe, Y.Nishimura and M.Matsubara, "Decomposition in large system optimization using the method of multipliers," J. Optimiz. Theory Appl., Vol. 25 (1978) 181-193.