A Multiplier Method with Variable Augmented Lagrangian Functions

新見　朋広

（京都大学大学院情報学研究科数理工学専攻　現所属：日本銀行）

指導教員：山下信雄 准教授

1. はじめに

本論文では，以下の非線形計画問題を考える．

$$\begin{array}{lll} \min_x & f(x) & \\ \text{s.t.} & g_i(x) \le 0 \ (i=1,\dots,l), & (1) \\ & h_j(x) = 0 \ (j=1,\dots,m) & \end{array}$$

ただし，関数 $f:\mathbb{R}^n \mapsto \mathbb{R}$, $g_i:\mathbb{R}^n \mapsto \mathbb{R}$, および $h_j:\mathbb{R}^n \mapsto \mathbb{R}$ は 2 回連続的微分可能とする．

この問題が大規模問題であるとき，内点法や逐次二次計画法（SQP 法）等は有効でなく，近年 Accelerated Proximal Gradient Method（APG 法）[1] 等の種々の射影勾配法が研究されている．しかし，射影勾配法は複雑な制約に適さないという問題がある．そのため，複雑な制約の問題に対しては，古典的手法である乗数法が用いられる．乗数法はラグランジュ関数に制約に対するペナルティ項を付加した拡張ラグランジュ関数 Q を導入し，Q を無制約あるいは簡単な制約の下で最小化して元の問題の局所最適解を求める手法である．乗数法の効率性は，Q 内のラグランジュ乗数に依存する．これが KKT 条件におけるラグランジュ乗数であれば安定的に解を求められる．そこで乗数法ではラグランジュ乗数を適宜更新して推定し，推定されたラグランジュ乗数を用いた Q の最小化を繰り返し行う．そのため，乗数法ではこの推定を効率的に行うことが重要となる．しかし，従来の乗数法では Q に組み込む制約はすべての制約，または固定された一部の制約であることが前提となっているため，推定に時間がかかることがある．

そこで本論文では，各反復で拡張ラグランジュ関数に組み込む制約を固定せずに交換可能とし，各反復では一部の制約を残した部分問題を解く乗数法を考える．これにより，ラグランジュ乗数の推定を効率化するとともに数値的不安定性を回避する手法を提案する．

2. 提案アルゴリズム

各反復で任意の一部の制約のみ組み込む拡張ラグランジュ関数を考え，残りの制約（残留制約）は制約に残した問題を部分問題として考える．k 回目の反復での残留不等式制約，残留等式制約の添え字集合をおのおの G_k, H_k とする．このとき拡張ラグランジュ関数 Q_k は $Q_k(x,\lambda,\mu,t,r) := f(x) + \sum_{i\notin G_k} \frac{1}{2t_i}\left[\max\{0,\lambda_i+t_i g_i(x)\}^2 - {\lambda_i}^2\right] + \sum_{j\notin H_k}\left[\mu_j h_j(x) + \frac{1}{2}r_j\{h_j(x)\}^2\right]$ と定義される．よって，毎回の反復で解くべき提案手法の部分問題は以下のようになる．

$$\begin{array}{ll} \min\limits_x & Q_k(x,\lambda^k,\mu^k,t^k,r^k) \\ \text{s.t.} & g_i(x)\le 0 \ (i\in G_k), \quad h_j(x)=0 \ (j\in H_k) \end{array} \qquad (2)$$

毎回の制約での残留制約 G_k, H_k を固定すれば提案手法は従来の乗数法と一致するため，提案手法は従来の乗数法の一般化手法とみなすことができる．

ラグランジュ乗数の推定

Q_k に組み込まれる制約 g_i, h_j $(i\notin G_k, j\notin H_k)$ に対するラグランジュ乗数については既存の推定式を用いる．しかし，残留制約 g_i, h_j $(i\in G_k, j\in H_k)$ に対するラグランジュ乗数については，部分問題 (2) を解いて得られたラグランジュ乗数 $\hat{\lambda}_i^k, \hat{\mu}_j^k$ を用いて更新することとする．すなわちラグランジュ乗数の更新式は以下のようになる．

$$\begin{aligned} \lambda_i^{k+1} &= \hat{\lambda}_i^k \ (i\in G_k)\ ; \lambda_i^{k+1} = \max\{0, \lambda_i^k + t_i^k g_i(x^k)\} \\ &\qquad (i\in\{1,\dots,l\}\setminus G_k), \\ \mu_j^{k+1} &= \hat{\mu}_j^k \ (j\in H_k); \mu_j^{k+1} = \mu_j^k + r_j^k h_j(x^k) \\ &\qquad (j\in\{1,\dots,m\}\setminus H_k). \end{aligned}$$

ただし，x^k は部分問題 (2) を解いて得られた解であり，t_i^k, r_j^k は各制約に対するペナルティパラメータである．

残留制約に対するラグランジュ乗数 $\hat{\lambda}_i^k, \hat{\mu}_j^k$ は元の問題 (1) における精度の良いラグランジュ乗数になることが期待できるため，毎回の反復で残留制約を交換していけば，残留制約を固定する従来の乗数法に比べてより効率よくラグランジュ乗数の推定ができる．

提案手法の収束性

提案手法の大域的収束性が，以下の定理で示される．

定理 1. 部分問題 (2) の解 x^k および対応するラグランジュ乗数 $\hat{\lambda}^k, \hat{\mu}^k$ は計算可能とする．さらに，ある部分列 $\{x^k\}_K$ が $\bar{x}$ に収束し，$\nabla g_i(\bar{x})$ $(i \in \{i \mid g_i(\bar{x}) = 0\}), \nabla h_j(\bar{x})$ $(j = 1, \ldots, m)$ は線形独立とする．このとき，部分列 $\{\lambda^k\}_K$, $\{\mu^k\}_K$ はおのおのある点 $\bar{\lambda}$, $\bar{\mu}$ に収束し，$(\bar{x}, \bar{\lambda}, \bar{\mu})$ は元の問題 (1) の KKT 条件を満たす．

なお上記の収束性は残留制約 G_k, H_k の選び方に依存しないため，残留制約の選び方を工夫することで収束を効率化することができる．論文中では，有効な残留制約の選び方の例を提案すとともに，部分問題の解に誤差を許容し，毎回の反復の部分問題を必ずしも正確に解かなくても大域的収束性が成り立つことを示している．

部分問題を解くアルゴリズム

提案手法の部分問題 (2) の解き方は重要である．本論文では APG 法等の射影勾配法を用いることを想定している．射影勾配法は，特定の簡単な制約（優位制約）を持つ問題を簡単に解くことができるという特徴を持つ [1]．そこで，部分問題の制約を任意に選択できる提案手法の利点を活かし，簡単な制約（特に優位制約）を残留制約に選べば部分問題は射影勾配法で簡単に解ける．そのため射影勾配法は提案手法の部分問題を解く手法として相性がよく，さらに射影勾配法が苦手とする複雑な制約の問題に対しても適用することができる．

3. 数値実験

機械学習の ν-SVM の双対問題に現れる大規模な凸計画問題を用いて数値実験を行った．従来の乗数法 (CMM) と提案手法 (SCM) を比較した結果が図 1，2 である．ϵ_1^k および ϵ_2^k は k 回目の反復における KKT 条件の誤差 $\epsilon_1^k := \frac{1}{n} \left\| \nabla f(x^k) + \sum_{j=1}^m \mu_j^k \nabla h_j(x^k) + \sum_{i=1}^l \lambda_i^k \nabla g_i(x^k) \right\|$, $\epsilon_2^k := \frac{1}{n} \sum_{j=1}^m \left| h_j(x^k) \right| + \frac{1}{n} \sum_{i=1}^l \left| \max\{g_i(x^k), -\lambda_i^k\} \right|$ である．ただし x^k は (2) を解いて得た解，μ_j^k, λ_i^k は k 回目の反復で推定されたラグランジュ乗数である．ϵ_3^k は，k 回目の反復でのラグランジュ乗数の推定誤差である．図 1 より，提案手法は従来の乗数法よりも効率的に局所的最適解に収束することがわかる．これは，図 2 からわかるように，部分問題で得たラグランジュ乗数を利用することで，KKT 条件を満たすラグランジュ乗数を効率よく推定できているためである．また論文中では，残留制約の選択方法を工夫することでペナルティパラメータが発散しにくくできることを確認した．

4. まとめ

本論文では，固定された拡張ラグランジュ関数を用いる従来の乗数法を改め，各反復で拡張ラグランジュ関数に組み込む制約を任意に選択し，部分問題で得たラグランジュ乗数をラグランジュ乗数の推定に利用する新たな乗数法を提案し，大域的収束性を示した．また，提案手法によりラグランジュ乗数の推定精度が向上すると共にペナルティパラメータの発散を防ぎ，数値的不安定性が起きにくくできることを確認した．

参考文献

[1] P. Tseng, Approximation accuracy gradient methods, and error bound for structured convex optimization, *Mathematical Programming*, Ser. B, **125**, 263–295, 2010.

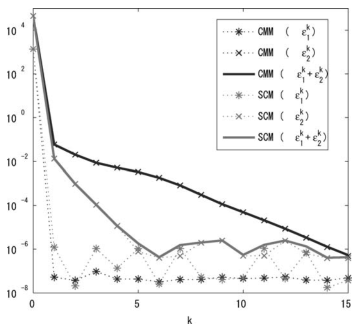

図 1　KKT 条件の精度比較

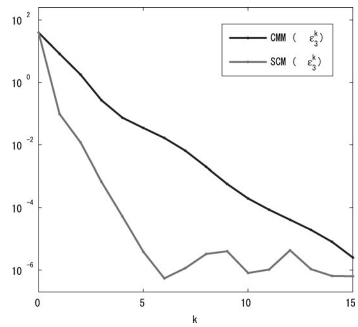

図 2　推定ラグランジュ乗数の精度比較