

高速な構造保存解法に基づく 連続最適化アルゴリズムの構成と解析

数理情報学専攻 48-196205 大沼 賢矢

指導教員 松尾 宇泰 教授

1 序論

常微分方程式と、無制約連続最適化問題に対する反復法の関連についての研究は古くから存在するが、その中でも勾配流と呼ばれる常微分方程式と連続最適化の関連を指摘する研究は多くある。また、勾配流はそれ自体物理現象のモデリングなどで重要な対象であり、その数値解法についても盛んに研究されている。特に勾配流の持つ散逸性と呼ばれる性質が離散化後も保たれるような数値解法は安定性などの面で有利であると考えられている。そのなかで、Cheng et al. (2020) により [1] で提案された未定乗数法アプローチは非常に高速であるが、反復の中で解かなければならない非線形方程式の可解性が未解決であるという問題がある。本論文では未定乗数法アプローチの特殊なケースについて可解性を証明する。また、未定乗数法アプローチを最適化手法とみなしたときの収束レートの解析や、最適化手法とみなす際の高速化について提案する。

2 勾配流に対する未定乗数法アプローチ

勾配流とは、

$$\dot{x} = -A\nabla f(x), \quad x(0) = x_0 \quad (1)$$

の形で書かれる常微分方程式である。ここで $n \in \mathbb{N}, x : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ は L -平滑かつ強圧的な関数であり、 $A \in \mathbb{R}^{n \times n}$ は正定値行列である。勾配流 (1) の重要な性質として散逸性

$$\frac{d}{dt} f(x) \leq 0$$

があげられる。散逸性を保つ数値解法として離散勾配法があるが、離散勾配法は陰的な手法となり計算量が大きくなってしまいう問題がある。そこでまず [1] では、対称行列 $Q \in \mathbb{R}^{n \times n}$ と $g : \mathbb{R}^n \rightarrow \mathbb{R}$ を用いて $f(x) = \langle x, Qx \rangle / 2 + g(x)$ と分解し、勾配流を補助変数 $\eta : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ を導入した等価な ODE

$$\dot{x} = -A(Qx + \nabla g(x)), \quad (2)$$

$$\frac{d}{dt} g(x) = \eta \langle \nabla g(x), \dot{x} \rangle \quad (3)$$

と変形し、未定乗数法アプローチ

$$\frac{x_{k+1} - x_k}{h} = -A \left(Q \left(\frac{x_{k+1} + x_k}{2} \right) + \eta_k \nabla g(x_{k+1/2}^*) \right)$$

$$g(x_{k+1}) - g(x_k) = \eta_k \langle \nabla g(x_{k+1/2}^*), x_{k+1} - x_k \rangle$$

を提案している。ここで $h > 0$ は時間刻み幅、 $x_k \approx x(kh)$ は数値解であり、 $x_{k+1/2}^*$ は x_{k+1} によらず、 $x((k+1/2)h)$ の近似となるベクトルである。この手法は離散散逸性

$$f(x_{k+1}) \leq f(x_k)$$

を満たす。未定乗数法アプローチにおいて x_k から x_{k+1} を計算する際は、2つの線形方程式とスカラー η_k の非線形方程式を解けばよく、高速である。ただし、非線形方程式の可解性は未解決である。

3 未定乗数法アプローチの特殊な場合の可解性

本節では未定乗数法アプローチで $Q = 0, x_{k+1/2}^* = x_k$ とした特殊な場合である

$$\frac{x_{k+1} - x_k}{h} = -\eta_k A \nabla f(x_k) \quad (4)$$

$$f(x_{k+1}) - f(x_k) = \eta_k \langle \nabla f(x_k), x_{k+1} - x_k \rangle \quad (5)$$

について考える。この場合は、 η_k の非線形方程式

$$F(\eta_k; x_k) := f(x_k - \eta_k h A \nabla f(x_k)) - f(x_k) + h \eta_k^2 \langle \nabla f(x_k), A \nabla f(x_k) \rangle = 0$$

を解けば x_{k+1} を求められる。この非線形方程式の可解性について次のことが言える。

定理 1. 任意の $x_k \in \mathbb{R}^n$ について、非線形方程式 $F(\eta_k; x_k) = 0$ は、 $\eta_k \geq (Lh/2\lambda_{\min} + 1)^{-1}$ を満たす解をもつ。ただし、 $\lambda_{\min}$ は正定値対称行列 $(A^{-1} + (A^{-1})^T)/2$ の最小固有値である。また、 f が凸でありかつ $\nabla f(x_k) \neq 0$ である場合、 $F(\eta_k; x_k) = 0$ の $\eta_k \neq 0$ なる解はただ一つである。

4 未定乗数法アプローチの収束レート

A が単位行列である場合の未定乗数法アプローチ (4), (5) を最適化手法としてみなす場合、次のような収束レートが保証される。

定理 2. 式 (4), (5) により定義される点列 $\{x_k\}_{k=0}^\infty$ は,

$$\min_{0 \leq i \leq k} \|\nabla f(x_i)\| \leq \left(\frac{Lh}{2} + 1\right) \sqrt{\frac{f(x_0) - f^*}{(k+1)h}}$$

を満たす. さらに f が凸である場合は

$$f(x_k) - f^* = O\left(\frac{1}{k}\right)$$

を満たし, さらに f が μ -強凸である場合は

$$f(x_k) - f^* \leq \exp\left(-\frac{8\mu kh}{(Lh+2)^2}\right)(f(x_0) - f^*)$$

を満たす.

また, 未定乗数法アプローチの高速化を考える. 離散散逸則が満たされるためには $F(\eta_k; x_k) \leq 0$ であればよく非線形方程式の解を求める必要がないこと, また最適化手法として用いる際には解軌道を精度よく求める必要がないことを考慮すると, 次のアルゴリズムのような最適化が可能である. (ここで, 勾配における許容誤差を $\epsilon > 0$ とする. ($\|\nabla f(x)\| < \epsilon$ となる x を求める.))

アルゴリズム 1 高速化した提案手法

Input: $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$,

$h > 0$, $\epsilon > 0$

$k \leftarrow 0$

while $\|\nabla f(x_k)\| \geq \epsilon$ **do**

$\eta_k \leftarrow 1$

while $F(\eta_k; x_k) > 0$ **do**

$\eta_k \leftarrow \frac{\eta_k}{2}$

end while

$x_{k+1} \leftarrow x_k - h\eta_k \nabla f(x_k)$

$k \leftarrow k + 1$

end while

$x \leftarrow x_k$

Output: $x \in \mathbb{R}^n$

この高速化したアルゴリズムも高速化する前の手法と同様, 次のような収束レートが保証される.

定理 3. 式 (4), (5) により定義される点列 $\{x_k\}_{k=0}^\infty$ は,

$$\min_{0 \leq i \leq k} \|\nabla f(x_i)\| \leq (Lh+2) \sqrt{\frac{f(x_0) - f^*}{(k+1)h}}$$

を満たす. さらに f が凸である場合は

$$f(x_k) - f^* = O\left(\frac{1}{k}\right)$$

を満たし, さらに f が μ -強凸である場合は

$$f(x_k) - f^* \leq \exp\left(-\frac{2\mu kh}{(Lh+2)^2}\right)(f(x_0) - f^*)$$

を満たす.

5 数値実験

目的関数が 2 次関数 $f(x) = \langle x, Ax \rangle / 2 + \langle b, x \rangle$ の場合に対して数値実験を行った結果の抜粋が図 1 である. なお高速化していない未定乗数法アプローチを提案手法 1, 高速化したものを提案手法 2 としている.

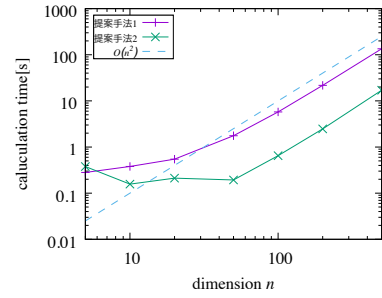


図 1. $h = 1$ としたときの提案手法の実行時間の比較

6 結論

勾配流に対する未定乗数法アプローチについて, これまで未解決であった非線形方程式の可解性を, 特殊なケースについて証明し, さらに目的関数が凸な場合の非自明な解の唯一性を証明した. また, 未定乗数法アプローチの特殊なケースを最適化手法とみなした時の収束レートの証明を行った. それに加え, 未定乗数法アプローチを最適化手法としてみる際の高速化手法を提案した. これらの結果は, パラメータが適応的に定められる最適化手法を常微分方程式の数値解法から自然に導出して, 連続版の収束レートとの対応も自然に確認できるという点で意義があると考えられる.

今後の課題としては, 未定乗数法アプローチの一般の場合の可解性の証明や, 提案手法のパラメータ h をどのように選択すれば収束性がよくなるかの検討があげられる. また, 他の常微分方程式にも同様の枠組みを適用してよりよい最適化手法を生み出せるかという課題もあげられる.

参考文献

- [1] Q. Cheng, C. Liu, and J. Shen. A New Lagrange Multiplier Approach for Gradient Flows. *Computer Methods in Applied Mechanics and Engineering*, 367:113070, 2020.