日本オペレーションズ・リサーチ学会
2021年 春季研究発表会

リーマン多様体上の共役勾配法の新しいクラスについて

05000081 京都大学 *佐藤寛之 SATO Hiroyuki

1. はじめに

連続最適化手法の中でも一次の情報のみを用いる一次法の一つに共役勾配法がある．ユークリッド空間上の最適化問題に対する共役勾配法は現在でも研究が続けられているが，それらをリーマン多様体上の手法に拡張する研究も重要である．

本研究では，リーマン多様体 $(M, \langle\cdot,\cdot\rangle)$ 上の関数 $f\colon M \to \mathbb{R}$ の最小化問題に対する共役勾配法の新たなクラスについて議論する．多様体 M にはレトラクション $R\colon TM \to M$ が与えられており，目的関数 f は下に有界かつ，ある $L > 0$ が存在し，任意の $t \geq 0, x \in M$, および，$\|\eta\|_x = 1$ なる $\eta \in T_xM$ について，$|\mathrm{D}(f \circ R_x)(t\eta)[\eta] - \mathrm{D}(f \circ R_x)(0)[\eta]| \leq Lt$ が成り立つと仮定する．ここで，T_xM は M 上の点 x における接空間，TM は M の接束を表す．また，リーマン計量 $\langle\cdot,\cdot\rangle$ から定まる T_xM の内積を $\langle\cdot,\cdot\rangle_x$ と書き，$\xi \in T_xM$ のノルムを $\|\xi\|_x := \sqrt{\langle\xi,\xi\rangle_x}$ と定義する．さらに，f の M 上の点 x におけるリーマン計量から定まる勾配を $\operatorname{grad} f(x)$ と書く．表記を簡単にするため，M 上の点列 $\{x_k\}$ に対して $g_k := \operatorname{grad} f(x_k)$ と書くことにする．

2. リーマン多様体上の種々の共役勾配法

ユークリッド空間 $\mathbb{R}^n$ での直線探索法で用いられる更新式 $x_{k+1} := x_k + t_k\eta_k$ を多様体に拡張する際には，レトラクション R と呼ばれる写像が重要となる [1]．具体的には，多様体 M 上で点 x_k および探索方向 $\eta_k \in T_{x_k}M$ が与えられたとき，探索を行うためには x_k から η_k の方向に伸びる M 上の曲線 γ が必要となる．そこで，レトラクションを用いて $\gamma(t) := R_{x_k}(t\eta_k)$ $(t \geq 0)$ と定義すると，探索の結果として得られたステップ幅 $t_k > 0$ を用いて

$$x_{k+1} := R_{x_k}(t_k\eta_k)$$

によって M 上の点列 $\{x_k\}$ を生成できる．

ユークリッド空間 $\mathbb{R}^n$ 上の共役勾配法での点 x_{k+1} における探索方向は $\eta_{k+1} = -g_k + \beta_{k+1}\eta_k$ と計算される．ここで，$\beta_{k+1} \in \mathbb{R}$ であり，$\eta_k \in \mathbb{R}^n$ は点 $x_k \in \mathbb{R}^n$ における探索方向である．一方，多様体上の共役勾配法においては，探索方向を計算する際に注意を要する．すなわち，探索方向 η_{k+1} は $T_{x_{k+1}}M$ の元として計算する必要があるが，$-g_{k+1} \in T_{x_{k+1}}M$ と $\beta_{k+1}\eta_k \in T_{x_k}M$ は異なる接空間に属するため足し合わせることができない．

この問題を解決する方法として，Smith [6] は，x_k から x_{k+1} への測地線に沿っての平行移動 P を用いて，$\eta_{k+1} := -g_{k+1} + \beta_{k+1}P(\eta_k)$ として探索方向を計算することを提案した．しかしながら，平行移動は計算量が大きくなることや，陽には計算できないこともある．そこで，Absil ら [1] は，平行移動の一般化として vector transport と呼ばれる写像 $\mathcal{T}\colon TM \oplus TM \to TM$ を定義し，探索方向を $\eta_{k+1} := -g_{k+1} + \beta_{k+1}\mathcal{T}_{t_k\eta_k}(\eta_k)$ と計算する共役勾配法の枠組みを提案した．

リーマン多様体上の共役勾配法のこの枠組みを用いることで，Ring & Wirth [2] は，ユークリッド空間における Fletcher–Reeves の β_{k+1} を拡張して $\beta^{\text{R-FR}}_{k+1} := \|g_{k+1}\|^2_{x_{k+1}}/\|g_k\|^2_{x_k}$ を提案した．ただし，こうして得られる共役勾配法の大域的収束性を保証するには条件

$$\|\mathcal{T}_{t_k\eta_k}(\eta_k)\|_{x_{k+1}} \leq \|\eta_k\|_{x_k} \tag{1}$$

を仮定する必要がある．

しかしながら，実際の最適化計算においては条件 (1) がしばしば満たされないという問題がある．そこで，条件 (1) を仮定する必要がないアルゴリズムを構築するために，Sato & Iwai [5] は scaled vector transport $\mathcal{T}^0$ を $\xi, \eta \in T_xM$ に対して $\mathcal{T}^0_\eta(\xi) := (\|\xi\|_x/\|\mathcal{T}^R_\eta(\xi)\|_{R_x(\eta)})\mathcal{T}^R_\eta(\xi)$ と定義した．なお，$\mathcal{T}^R_\eta(\xi) := \mathrm{D}R_x(\eta)[\xi]$ はレトラクションの微分によって得られる vector transport である．そして，条件 (1) が成り立たない場合に限り $\mathcal{T}^0$ を用い，(1) が成り立つときには $\mathcal{T}^R$ をそのまま用いて探索方向を計算する．この改善により，$\beta^{\text{R-FR}}_{k+1}$ とウルフ条件を満たすステップ幅を用いた共役勾配法が大域的収束性をもつことが，(1) を仮定せずに証明される．実は，scaled vector transport $\mathcal{T}^0_\eta(\xi)$ は ξ についての線形性をもたず vector transport の

定義を満たさないことが，本研究では重要となる．

また，Sato [4] はユークリッド空間での Dai–Yuan の $\beta_{k+1}^{\mathrm{DY}}$ をリーマン多様体上に $\beta_{k+1}^{\mathrm{R\text{-}DY}} := \|g_{k+1}\|_{x_{k+1}}^2/(\langle g_{k+1}, \mathcal{T}_{t_k\eta_k}^{(k)}(\eta_k)\rangle_{x_{k+1}} - \langle g_k, \eta_k\rangle_{x_k})$ と拡張し，結果として得られる共役勾配法の大域的収束性を証明した．さらに，Sakai & Iiduka [3] は，$-\sigma \le r_k \le 1$ を満たす β_k を用いるリーマン多様体上の共役勾配法のクラスに対して大域的収束性の証明を与えた．ここで，$r_k := \beta_k/\beta_k^{\mathrm{R\text{-}DY}}$, $\sigma := (1-c_2)/(1+c_2)$ であり，c_2 はステップ幅に対する強ウルフ条件に現れる定数である．

3. リーマン多様体上の共役勾配法の新しいクラス

これまでのリーマン多様体上の共役勾配法の研究では，β_{k+1} の計算方法の提案と，その β_{k+1} を用いた場合の大域的収束性の解析が主眼に置かれることが多かったが，最近の Zhu & Sato による研究 [7] では，vector transport $\mathcal{T}$ を用いない共役勾配法が提案された．すなわち，M 上の 2 つのレトラクション R^{fw} および R^{bw}（これらは等しくても良い）を与えておき，$x_{k+1} := R_{x_k}^{\mathrm{fw}}(t_k\eta_k)$ かつ，$\eta_{k+1} := -g_{k+1} - \beta_{k+1}s_k t_k^{-1}(R_{x_{k+1}}^{\mathrm{bw}})^{-1}(x_k)$ によって点列 $\{x_k\}$ を生成してゆく．ここで，正の実数 $s_k := \min\left\{1, \|\eta_k\|_{x_k}/\|t_k^{-1}\left(R_{x_{k+1}}^{\mathrm{bw}}\right)^{-1}(x_k)\|_{x_{k+1}}\right\}$ は前節の scaled vector transport と同じく，大域的収束性を保証するためのスケーリングの役割を果たす．

以上から，リーマン多様体上の共役勾配法について以下のように考察することができる．まず，平行移動を用いることは理論的には自然であるが，実際の計算ではしばしば困難を伴う．一方で，平行移動を一般化した vector transport を用いる共役勾配法は，vector transport が条件 (1) を満たさないという点で一般的すぎて失敗することがあると同時に，vector transport の線形性が手法を特殊なものに制限してしまい得るという側面もある．共役勾配法の収束性の観点からは，vector transport を改善したものとして scaled vector transport を位置付けることができるし，vector transport の代替として $\left(R^{\mathrm{bw}}\right)^{-1}$ を利用する方法を用いることもできる．

これらをすべて統合することで，リーマン多様体上の共役勾配法の新たな枠組みとして，探索方向を

$$\eta_{k+1} := -g_{k+1} + \beta_{k+1}\mathcal{T}^{(k)}(\eta_k)$$

と計算するアルゴリズムを提案する．ここで，$\mathcal{T}^{(k)}\colon T_{x_k}M \to T_{x_{k+1}}M$ は vector transport とは限らない一般的な写像で，(1) に対応する条件 $\|\mathcal{T}^{(k)}(\eta_k)\|_{x_{k+1}} \le \|\eta_k\|_{x_k}$ を満たすものとする．これは [6, 5, 7] の枠組みをその特殊例として含んでおり，上記の仮定は提案手法の実用性を損なわない．当日の講演では，提案手法の一般性，あるクラスの β_k についての収束性，および，具体例を用いての数値実験による検証について議論する．

参考文献

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, 2008.

[2] W. Ring and B. Wirth, Optimization methods on Riemannian manifolds and their application to shape space, *SIAM Journal on Optimization*, **22**(2), 596–627, 2012.

[3] H. Sakai and H. Iiduka, Hybrid Riemannian conjugate gradient methods with global convergence properties, *Computational Optimization and Applications*, **77**(3), 811–830, 2020.

[4] H. Sato, A Dai–Yuan-type Riemannian conjugate gradient method with the weak Wolfe conditions, *Computational Optimization and Applications*, **64**(1), 101–118, 2016.

[5] H. Sato and T. Iwai, A new, globally convergent Riemannian conjugate gradient method, *Optimization*, **64**(4), 1011–1031, 2015.

[6] S. T. Smith, Optimization techniques on Riemannian manifolds, *Hamiltonian and Gradient Flows, Algorithms and Control*, 113–135, American Mathematical Society, 1994.

[7] X. Zhu and H. Sato, Riemannian conjugate gradient methods with inverse retraction, *Computational Optimization and Applications*, **77**(3), 779–810, 2020.