

一般化シュティーフェル多様体上のレトラクション とその効果的な実装について

東京理科大学・工学部情報工学科 佐藤 寛之

東京理科大学・理学部数理情報科学科 相原 研輔

Hiroyuki Sato¹

Department of Information and Computer Technology, Tokyo University of Science

Kensuke Aihara

Department of Mathematical Information Science, Tokyo University of Science

概要

近年、ユークリッド空間における最適化アルゴリズムをリーマン多様体上に拡張する研究が盛んに行われており、様々な分野への応用が期待されている。リーマン多様体上の最適化問題に対する反復法では、各点における探索方向をその点での接ベクトルとして与え、探索方向に進んだ点を多様体上に写すことで点列が更新される。この操作を実現する写像をレトラクションと呼び、シュティーフェル多様体上の最適化では行列のQR分解に基づくレトラクションがよく用いられる。本稿ではまず、このQR分解に基づくレトラクションを一般化シュティーフェル多様体上に拡張する。次に、拡張したレトラクションについて、コレスキーQR分解を利用した効果的なレトラクションの計算法を提案する。そして、提案する方法の有効性を数値実験により検証する。

1 はじめに

ある領域において定義された関数を最小化する問題を最適化問題といい、最小化しようとする関数を目的関数と呼ぶ。ユークリッド空間における無制約の最適化問題の場合、目的関数の勾配やヘッセ行列を用いることで、ニュートン法や共役勾配法などの効果的な反復法を適用することができる。一方、制約条件付きの最適化問題の場合、ニュートン法などの通常の最適化アルゴリズムは、生成される点列が制約条件を満たさないため、一般には適用することができない。しかし、制約を満たす実行可能領域がリーマン多様体をなす場合には、対象とする問題をリーマン多様体上の無制約の最適化問題として再定式化することができる [1]。そこで、ユークリッド空間上の最適化アルゴリズムをリーマン多様体上に拡張した新しい最適化手法の研究が盛んに行われている [2, 7]。最近では、正準相関分析などの統計手法や制御理論におけるモデル低次元化などの問題もリーマン多様体上の最適化問題として定式化できることが知られており [8, 9]、多様体上の最適化手法の幅広い応用が期待されている。

¹e-mail: [hsato\[AT\]rs.tus.ac.jp](mailto:hsato[at]rs.tus.ac.jp)

本稿では、リーマン多様体上の反復法において重要な、点列の更新方法について議論する。一般に、点列を更新する際には、各点における探索方向をその点での接ベクトルとして与え、探索方向に進んだ点を多様体上に写すレトラクションと呼ばれる写像が必要となる。具体的に、シュティーフェル多様体上の最適化では行列の極分解やQR分解に基づくレトラクションがよく用いられる [1, 7]。一方、一般化シュティーフェル多様体上の最適化については、文献 [9] において極分解に基づくレトラクションが導入されているが、QR分解に基づく方法はこれまでほとんど議論されていない。そこで本稿では、QR分解に基づくシュティーフェル多様体上のレトラクションを一般化シュティーフェル多様体上に拡張する。このとき、通常のQR分解に加えて、行列の平方根やその逆行列の計算が現れるため、問題の規模が大きくなるにつれて計算コストが増大し、実用的ではなくなる。そこで我々は、コレスキーQR分解 [4] を利用した計算効率の良いレトラクションの実装法を新たに提案する。数値実験を通して、提案するレトラクションは極分解に基づく方法に比べて短い計算時間で実行できることを示す。

本稿の構成は以下の通りである。まず2節で、リーマン多様体上の最適化の考え方と一般的なレトラクションの定義について述べる。次に3節では、シュティーフェル多様体上のレトラクションとして、極分解とQR分解に基づく2種類の方法を紹介する。4節で、QR分解に基づく一般化シュティーフェル多様体上のレトラクションを示すとともに、コレスキーQR分解を利用した効率的な計算法を提案する。5節で、極分解に基づく既存のレトラクションとコレスキーQR分解に基づく方法を数値実験により比較し、提案手法の有効性を示す。最後に、6節でまとめを行う。

2 リーマン多様体上の最適化とレトラクション

目的関数 f についてのリーマン多様体 M 上の最適化問題を考える。

$$\begin{aligned} & \text{minimize} && f(x), \\ & \text{subject to} && x \in M. \end{aligned} \tag{2.1}$$

ユークリッド空間上の無制約最適化問題に対する直線探索法では、点 x_k において探索方向 η_k を計算し、次の点を $x_{k+1} := x_k + t_k \eta_k$ と定める。ここで、 $t_k > 0$ はステップ幅であり、 η_k によって定まる半直線 $\{x_k + t\eta_k\}_{t \geq 0}$ 上で目的関数値が十分に小さくなるように与える。しかし、このような点列の更新は、リーマン多様体上の問題 (2.1) に対しては意味をなさない。なぜなら、点 $x_k \in M$ における探索方向 η_k は、 x_k の接空間 $T_{x_k}M$ 上の接ベクトルとして計算されるが、一般に $x_k + t_k \eta_k$ は M 上の点とはならないからである。そこで、点 $x_k + t_k \eta_k$ を M 上に引き戻す操作が必要となり、この操作を実現する写像をレトラクションと呼ぶ。

一般に、リーマン多様体 M において x_k が与えられているとき、次の点 x_{k+1} の探索は、 $\gamma(0) = x_k \in M$, $\dot{\gamma}(0) = \eta_k \in T_{x_k}M$ なる M 上の曲線 γ に沿って行われる。すなわち、 M 上のレトラクションとは、解を探索する上で妥当な曲線を定めるような接バンドル $TM := \cup_{x \in M} T_x M$ から多様体 M への写像 $R: TM \rightarrow M$ のことであり、そのような写

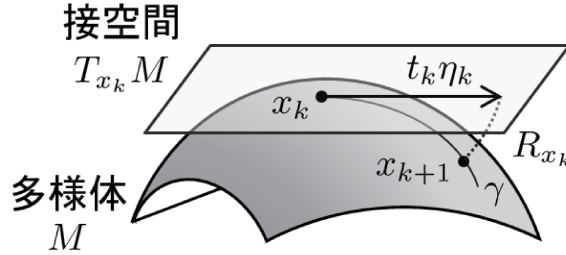


図 2.1: リーマン多様体上のレトラクション.

像が見つければ, 次の点を

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

として更新することができる (図 2.1 参照). ただし, $R_x := R|_{T_x M}$ は写像 R の $T_x M$ への制限である.

定義 2.1. M を多様体とし, TM を M の接バンドルとする. また, $R: TM \rightarrow M$ を滑らかな写像とし, R_x を R の $T_x M$ への制限とする. R が次の二つの条件を満たすとき, R を多様体 M 上のレトラクションという.

1. $T_x M$ の零ベクトル 0_x に対して, $R_x(0_x) = x$ が成り立つ.
2. $T_{0_x} T_x M \simeq T_x M$ という同一視の下で, R_x が

$$DR_x(0_x) = \text{id}_{T_x M}$$

を満たす. ここで, $DR_x(0_x)$ は写像 $R_x: T_x M \rightarrow M$ の 0_x における微分であり, $\text{id}_{T_x M}$ は $T_x M$ 上の恒等写像を表す.

3 シュティーフェル多様体上のレトラクション

$\text{St}(p, n)$ を $n \times p$ の列直交行列全体からなるシュティーフェル多様体

$$\text{St}(p, n) := \{Y \in \mathbb{R}^{n \times p} \mid Y^T Y = I_p\} \quad (3.1)$$

とし, 目的関数 F についての $\text{St}(p, n)$ 上の最適化問題を考える.

$$\begin{aligned} & \text{minimize} && F(U), \\ & \text{subject to} && U \in \text{St}(p, n). \end{aligned}$$

点 U_k と探索方向 ξ_k が与えられたとき, 前述のように $U_k + t_k \xi_k$ は一般に $\text{St}(p, n)$ 上の点とはならないため, これを $\text{St}(p, n)$ 上に引き戻すレトラクションが必要となる. この操

作は, U_k における接空間 $T_{U_k}\text{St}(p, n)$ から $\text{St}(p, n)$ への写像 $R_{U_k} : T_{U_k}\text{St}(p, n) \rightarrow \text{St}(p, n)$ を用いて次のように表される.

$$U_{k+1} := R_{U_k}(t_k \xi_k).$$

点 $U \in \text{St}(p, n)$ と接ベクトル $\xi \in T_U\text{St}(p, n)$ が与えられたとき, 具体的なレトラクションの計算法として, 極分解と QR 分解に基づく 2 種類の方法が知られている [1, 7]. 以降では, 簡単のためステップ幅を省略し, それぞれの計算法について概説する.

まず, 極分解に基づくレトラクションは, 次のように与えられる.

$$R_U(\xi) := (U + \xi) \sqrt{(I_p + \xi^T \xi)^{-1}}. \quad (3.2)$$

これは, $n \times p$ 列直交行列 Y と p 次半正定値対称行列 P を用いて $U + \xi = YP$ と極分解した場合における行列 $Y \in \text{St}(p, n)$ に相当している. ただし, ξ は点 $U \in \text{St}(p, n)$ における接ベクトルであることから, $U^T \xi + \xi^T U = 0$ が成り立ち [1],

$$(U + \xi)^T (U + \xi) = I_p + \xi^T \xi \quad (3.3)$$

と簡略化できることを用いている.

次に, QR 分解に基づくレトラクションを以下に示す.

$$R_U(\xi) := \text{qf}(U + \xi). \quad (3.4)$$

ここで, $\text{qf}(\cdot)$ は行列の QR 分解の Q 成分を返す関数である. QR 分解は, 高速かつ高精度な計算法が多く提案されており [4, 5], 式 (3.4) は実用的なレトラクションであると言える.

レトラクション (3.2) および (3.4) は, 定義 2.1 を満たすことが示せる [1]. 次節では, これらの一般化シュティーフェル多様体上への拡張について述べる.

4 一般化シュティーフェル多様体上のレトラクション

$G \in \mathbb{R}^{n \times n}$ を正定値対称行列とすると, 以下で与えられる多様体 $\text{St}_G(p, n)$ を一般化シュティーフェル多様体という.

$$\text{St}_G(p, n) := \{Y \in \mathbb{R}^{n \times p} \mid Y^T G Y = I_p\}.$$

これは, $G = I_n$ のとき, 通常のシュティーフェル多様体 (3.1) に一致する. 一方, $G \neq I_n$ の場合, (3.1) とは異なる直交条件の下での列直交行列の集合であることから, $\text{St}_G(p, n)$ 上で最適化を行うためには, レトラクションについても G に関して一般化する必要がある.

4.1 極分解と QR 分解に基づくレトラクションの拡張

$\text{St}_G(p, n)$ 上の点 U とその点における接ベクトル $\xi \in T_U\text{St}_G(p, n)$ が与えられているとする. このとき, 文献 [9] では, 極分解に基づくレトラクション (3.2) を拡張した次の方法が取り上げられている.

$$R_U^G(\xi) := (U + \xi) \sqrt{(I_p + \xi^T G \xi)^{-1}}. \quad (4.1)$$

これは, $Y^T G Y = I_p$ を満たす $n \times p$ 行列 Y と p 次半正定値対称行列 P を用いて $U + \xi = Y P$ と分解した場合における行列 $Y \in \text{St}_G(p, n)$ に相当する. ただし, (3.3) に対応して,

$$(U + \xi)^T G (U + \xi) = I_p + \xi^T G \xi \quad (4.2)$$

が成り立つことに注意する.

一方, QR 分解に基づく一般化シュティーフェル多様体上のレトラクションについては, その存在はほぼ明らかであるものの, これまで具体的な計算法についてはほとんど議論されていない. そこで本稿では, QR 分解に基づくレトラクションについて, まずはその陽的な表現を与える.

行列 $\sqrt{G}(U + \xi)$ に対する通常の QR 分解を考えることで, レトラクション (3.4) を拡張することができる. 具体的には, 次のように定める.

$$R_U^G(\xi) := \sqrt{G}^{-1} \text{qf}(\sqrt{G}(U + \xi)). \quad (4.3)$$

ここで, R^G は一般化シュティーフェル多様体 $\text{St}_G(p, n)$ 上のレトラクションとなることを定義 2.1 に沿って示す.

命題 4.1. (4.3) によって定義される写像 R^G は $\text{St}_G(p, n)$ 上のレトラクションである.

証明. $X \in \text{St}_G(p, n)$ とすると,

$$(\sqrt{G}X)^T (\sqrt{G}X) = X^T G X = I_p$$

が成り立つので, $\sqrt{G}X$ は $\text{St}(p, n)$ 上の点である. よって, $\sqrt{G}X = \sqrt{G}X \cdot I_p$ は $\sqrt{G}X$ の QR 分解であり, 分解の一意性から $\text{qf}(\sqrt{G}X) = \sqrt{G}X$ である. したがって,

$$R_X^G(0) = \sqrt{G}^{-1} \text{qf}(\sqrt{G}X) = \sqrt{G}^{-1} \sqrt{G}X = X$$

となり, 定義 2.1 の第 1 の条件が示された.

第 2 の条件については, 任意の $Y \in \text{St}(p, n)$ および $Z \in T_Y \text{St}(p, n)$ に対して $D \text{qf}(Y)[Z] = Z$ が成り立つことに注意する [1]. また, $\sqrt{G}X \in \text{St}(p, n)$ であり, $\xi \in T_X \text{St}_G(p, n)$ に対して

$$(\sqrt{G}\xi)^T \sqrt{G}X + (\sqrt{G}X)^T \sqrt{G}\xi = \xi^T G X + X^T G \xi = 0$$

が成り立つので, $\sqrt{G}\xi \in T_{\sqrt{G}X} \text{St}(p, n)$ である. これらを用いると, 任意の接ベクトル $\xi \in T_X \text{St}_G(p, n)$ に対して

$$DR_X^G(0)[\xi] = \sqrt{G}^{-1} D \text{qf}(\sqrt{G}X)[\sqrt{G}\xi] = \sqrt{G}^{-1} \sqrt{G}\xi = \xi$$

が成り立つことが示される. ξ は任意であるから, 上式は $DR_X^G(0) = \text{id}_{T_X \text{St}_G(p, n)}$ と同値である. ゆえに, 定義 2.1 の第 2 の条件も成り立ち, R^G が $\text{St}_G(p, n)$ 上のレトラクションであることが証明された.

なお, (4.3) は, $Q_G^\top G Q_G = I_p$ を満たす $n \times p$ 行列 Q_G と p 次上三角行列 R を用いて $U + \xi = Q_G R$ と分解した場合における Q_G に相当することを付記しておく. このことから, 標準内積を用いた QR 分解に基づくレトラクション (3.4) に対して, (4.3) は G を計量行列とする内積 (G -内積) を用いた場合への自然な拡張になっていると言える. また, (4.3) を G -内積を用いた QR 分解と再解釈すれば, 様々な計算法が考えられるが [6], 次節では, 最も効率的であると思われるコレスキー QR 分解に基づく方法を取り上げる.

4.2 コレスキー QR 分解に基づく効率的な実装

レトラクション (4.3) の実装法について考える. (4.3) では, 行列 G の平方根やその逆行列が現れるため, 大規模な問題に対してそのままの計算式を用いると, 多くの計算コストを要する. そこで本稿では, コレスキー QR 分解 [4] を利用した効率的な計算法を提案する.

一般に, フルランク行列 $A \in \mathbb{R}^{n \times p}$, $n \geq p$ のエコノミーサイズの QR 分解を $A = QR$ とするとき, 行列 R は $A^\top A$ のコレスキー分解によって得られる p 次三角行列に一致する [4]. このことを利用して, コレスキー QR 分解は,

$$A^\top A = R^\top R, \quad (4.4)$$

$$\text{qf}(A) = AR^{-1} \quad (4.5)$$

という二段階の計算によって行われる. ただし, (4.4) は $A^\top A$ のコレスキー分解を表す.

さて, 点 $U \in \text{St}_G(p, n)$ とその接ベクトル $\xi \in T_U \text{St}_G(p, n)$ が与えられたとき, $A := \sqrt{G}(U + \xi)$ として, 上記の手順 (4.4), (4.5) を適用する. ここで, (4.2) より,

$$A^\top A = I_p + \xi^\top G \xi$$

であるから, (4.4) は $I_p + \xi^\top G \xi$ のコレスキー分解を計算すればよい. そして,

$$\text{qf}(A) = AR^{-1} = \sqrt{G}(U + \xi)R^{-1}$$

であるから, レトラクションとして,

$$R_U^G(\xi) = \sqrt{G}^{-1} \text{qf}(A) = (U + \xi)R^{-1} \quad (4.6)$$

を求めればよいことが分かる.

以上より, (4.3) の実装として, まず $I_p + \xi^\top G \xi = R^\top R$ とコレスキー分解し, 得られた p 次上三角行列 R に対して, $(U + \xi)R^{-1}$ を求めればよい. 提案するレトラクションの計算アルゴリズムを, アルゴリズム 1 に示す.

一般に, n 次正定値対称行列の平方根やその逆行列の計算には少なくとも $O(n^3)$ の計算量を要するが, 提案手法では, \sqrt{G} や \sqrt{G}^{-1} を陽に求める必要はない. また, 実問題では $p \ll n$ である場合が多いため, (4.6) における p 次三角行列 R による逆変換の計算量は, n 次行列に関する演算に比べて相対的に小さいものと想定される. したがって, 提案手法における主要な計算は行列ベクトル積 $G\xi$ となり, 計算量は $O(pn^2)$ と評価できる. また, ア

アルゴリズム 1. コレスキー QR 分解に基づく $\text{St}_G(p, n)$ 上のレトラクション.

% 正定値対称行列 $G \in \mathbb{R}^{n \times n}$, $U \in \text{St}_G(p, n)$, $\xi \in T_U \text{St}_G(p, n)$ に対して以下を計算する.

1. $Z = I_p + \xi^\top G \xi$
 2. コレスキー分解 $Z = R^\top R$ を求める.
 3. $R_U^G(\xi) = (U + \xi)R^{-1}$
-

ルゴリズムの 1 行目については, 行列 $Z (= A^\top A)$ を $(U + \xi)^\top G (U + \xi)$ として素朴に計算する方法も考えられ, 一回のレトラクションとしては, こちらの方が計算量は少ない. しかし, 実用上は探索方向 ξ に対して $R_U^G(t\xi)$ における目的関数値が十分に下がるよう, ステップ幅 t を選択する必要がある. このような曲線上の探索においては, 異なる t に対して $(U + t\xi)^\top G (U + t\xi)$ を何度も計算するよりも, 行列ベクトル積 $G\xi$ を一度だけ計算して保持しておき, $I_p + t^2 \xi^\top G \xi$ を求める方法の方がはるかに効率が良いと言える. ただし, 丸め誤差の蓄積により (4.2) の等価性が大きく崩れる場合, すなわち ξ が点 U の接ベクトルとはならない場合は, $(U + t\xi)^\top G (U + t\xi)$ を素朴に計算する方法の方が丸め誤差の影響を抑えられると考えられる.

5 数値実験

本節では, 一般化シュティーフェル多様体上のレトラクションについて, 極分解に基づく方法と提案したコレスキー QR 分解に基づく方法を数値実験により比較し, 提案した方法の有効性を示す.

数値実験は, PC(Intel Core i7-3970X CPU 3.50GHz, 32.0GB RAM) において, MATLAB R2012a の倍精度演算を用いて行われた. レトラクションの比較のため, $k = 0, 1, 2, \dots$ に対して以下の計算を反復する.

1. $n \times p$ 乱数行列を U_k の接空間に射影して $\xi_k \in T_{U_k} \text{St}_G(p, n)$ を生成する. (5.1)
2. レトラクション R^G を用いて $U_{k+1} = R_{U_k}^G(\xi_k)$ を計算する. (5.2)

ただし, n 次正定値対称行列 G と初期点 $U_0 \in \text{St}_G(p, n)$ はそれぞれ乱数行列を元に生成する. 反復回数は 5 回とし, (5.1) を除いたレトラクションによる点の更新 (5.2) にかかる総計算時間を比較する. また, 5 反復終了時点での U_5 の直交性を示す指標として, $\|I_p - U_5^\top G U_5\|_F / \|I_p\|_F$ の値を調べる. 行列サイズは n を 3000 と固定した上で, p を 300, 600, 900, 1200, 1500 と変化させた.

提案したコレスキー QR 分解に基づくレトラクションの実装は, アルゴリズム 1 に従う. 極分解に基づくレトラクションは, 多様体上の最適化を行うためのツールボックス Manopt [3] を用いた. Manopt では, (4.1) の実装として, 行列 $(U + \xi)^\top G (U + \xi)$ の固有値分解に基づく方法が採用されている. 一回のレトラクションの計算量は $O(pn^2)$ と見積

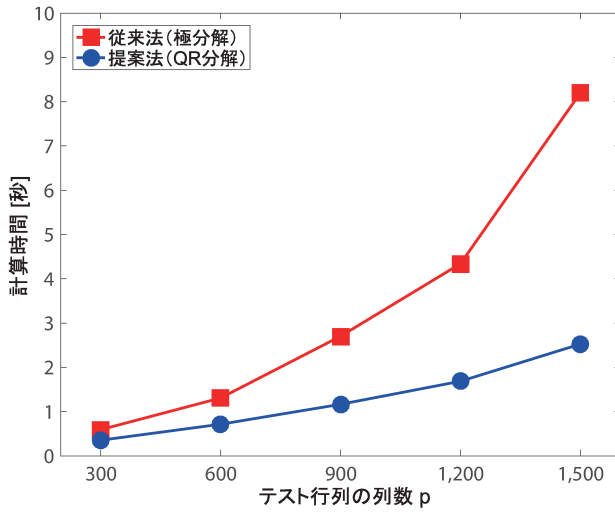


図 5.1: 極分解と QR 分解に基づくレトラクションの計算時間.

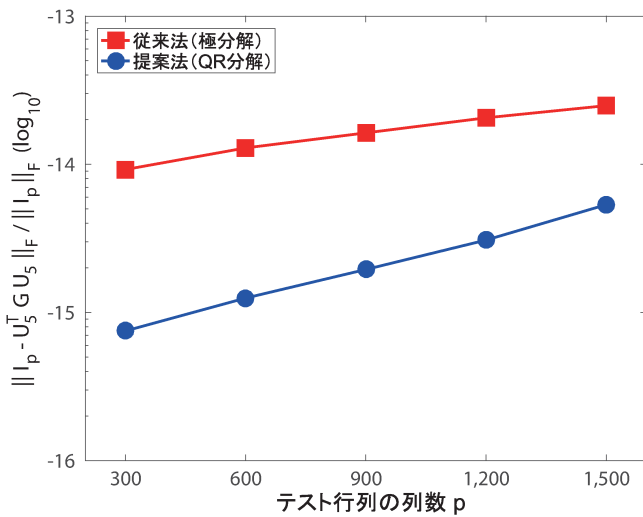


図 5.2: 極分解と QR 分解に基づくレトラクションの直交性.

もれるが、 p 次行列の固有値分解を用いるため、 p が大きくなると固有値分解に必要な計算時間が大幅に増加すると予想される。一方、提案した方法は p 次行列のコレスキー分解を用いるため、固有値分解に比べると計算量は少ないと言える。Manopt における (4.1) の実装の詳細は、公開されているソースコードを参照されたい。

レトラクションの総計算時間および反復終了時点での U_5 の直交性をそれぞれ図 5.1, 5.2

に示す。グラフの縦軸は、図 5.1 では計算時間 [秒]、図 5.2 では $\|I_p - U_5^T G U_5\|_F / \|I_p\|_F$ を対数表示した値を表し、横軸はいずれもテスト行列の列数 p を表す。

図 5.1, 5.2 より、次のことが言える。まず、計算時間については、極分解に基づく方法に比べて提案したコレスキー QR 分解に基づく方法の方が短い。特に、テスト行列の列数 p が大きくなるにつれて、その差は顕著であり、極分解に基づく方法では指数的に計算時間が増加しているが、提案した方法は概ね線形の増加である。この違いは、極分解に基づく方法が p 次行列の固有値分解を用いるのに対して、提案した方法はコレスキー分解を用いることに由来すると考えられる。次に、直交性については、いずれの方法でも $\|I_p - U_5^T G U_5\|_F / \|I_p\|_F$ の値が 10^{-13} を下回っており、この数値例では十分な直交性が得られているが、提案した方法の方が 1 桁程度高い直交性を維持している。また、行列の列数 p が大きくなることによる直交性の崩れは、いずれの方法でも僅かであった。

6 まとめ

本稿では、QR 分解に基づくシュティーフェル多様体上のレトラクションを、一般化シュティーフェル多様体上の場合に拡張するとともに、コレスキー QR 分解に基づく効率的な計算法を提案した。提案した方法は、一般化シュティーフェル多様体において導入される正定値対称行列 G の平方根やその逆行列の計算が不要であるという利点を持つ。また、数値実験では、Manopt にて実装されている極分解に基づくレトラクションに比べて短い計算時間で実行できることを示した。これは、 p 次行列の固有値分解を利用する Manopt 上の実装に対して、提案した方法はより計算量の少ないコレスキー分解を利用しているためであり、 p が大きいほどその効果も大きくなる。

直交性については、乱数行列を用いた模擬的な数値実験ではあるが、提案した方法において十分な精度の直交性が得られることを確認した。一般に、対象とする行列の条件数が大きいとき、コレスキー QR 分解は数値的に不安定となり、直交性が大きく崩れる場合がある [6]。しかし、一般化シュティーフェル多様体上の最適化においては、 $U^T G U = I_p$ を満たすように反復を行うため、最適解の近傍において $0 < t \ll 1$ となる場合、 $I_p + t^2 \xi^T G \xi$ は単位行列に近くなり、その条件数は大きくはならないであろう。このことと本稿の実験結果から、一般化シュティーフェル多様体上のレトラクションとして、コレスキー QR 分解に基づく方法は十分に実用に耐えうると期待される。

提案した方法の数値的安定性の解析、ならびに具体的な最適化手法への適用については、今後の課題としたい。

謝辞

本研究は、科学研究費補助金（若手研究 (B) : 16K17647, 15K17498）の助成を受けている。また、本研究について、貴重なご意見を頂いた今倉暁先生（筑波大学）、山本有作先生（電気通信大学）に深くお礼申し上げます。

参考文献

- [1] P.-A. Absil, R. Mahony, R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [2] K. Aihara, H. Sato, A matrix-free implementation of Riemannian Newton's method on the Stiefel manifold, *Optim. Lett.*, 2016, doi:10.1007/s11590-016-1090-9.
- [3] N. Boumal, B. Mishra, P.-A. Absil, R. Sepulchre, Manopt, a Matlab Toolbox for Optimization on Manifolds, *J. Mach. Learn. Res.*, **15** (2014), 1455–1459.
- [4] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 4th ed., The Johns Hopkins University Press, Baltimore, 2013.
- [5] N.J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
- [6] B.R. Lowery, J. Langou, Stability analysis of QR factorization in an oblique inner product, arXiv:1401.5171, 2014.
- [7] H. Sato, T. Iwai, A Riemannian optimization approach to the matrix singular value decomposition, *SIAM J. Optim.*, **23** (2013), 188–212.
- [8] W.-Y. Yan, J. Lam, An approximate approach to H^2 optimal model reduction, *IEEE Trans. Autom. Control*, **44** (1999), 1341–1358.
- [9] F. Yger, M. Berar, G. Gasso, A. Rakotomamonjy, Adaptive canonical correlation analysis based on matrix manifolds, *Proc. 29th International Conference on Machine Learning* (2012), 1071–1078.