

グラスマン多様体上の最適化アルゴリズム

京都大学大学院情報学研究科数理工学専攻

佐藤 寛之 (Hiroyuki Sato), 岩井 敏洋 (Toshihiro Iwai)

Department of Applied Mathematics and Physics,
Kyoto University

概要

制約条件付き最適化問題の実行可能領域がリーマン多様体をなす場合, その問題をユークリッド空間における制約条件付きの問題ではなく, リーマン多様体上の制約条件なしの問題であると見なすことができる. すると, 最急降下法やニュートン法などの制約条件なしの最適化手法をリーマン多様体上に拡張したアルゴリズムを適用することによってその問題が解ける. 本稿では, 最初にリーマン多様体上の最適化手法を概説する. その後, 行列の固有値問題への応用を見据え, グラスマン多様体上のレイリー商最小化問題を取り上げ, その幾何学および解法を紹介する.

1 序論

制約条件なしの最適化問題に対するアルゴリズムは, 制約条件付きの最適化手法より簡単であるものが多い. しかし, 制約条件付きの最適化問題であっても, その実行可能領域がリーマン多様体をなす場合は, その問題をリーマン多様体上の制約条件なしの最適化問題であると見なすことができる.

本稿ではまず, そうすることの利点を示すために, 多様体上の最適化問題と見なせる問題の例をいくつか紹介する. 次に, リーマン多様体上の最適化問題の一般論を紹介した後, 具体的にグラスマン多様体上の最適化アルゴリズムを導出する. 特に, グラスマン多様体上のレイリー商最適化問題について詳しく検討する.

2 リーマン多様体上の最適化問題の例

この節では, リーマン多様体上の最適化問題の例をいくつか紹介する.

2.1 レイリー商最小化問題と球面上の最適化問題

まず次の最適化問題を考えよう. ただし, A は n 次対称行列とする.

問題 2.1.

$$\begin{aligned} & \text{minimize } f(x) = \frac{x^T A x}{x^T x}, \\ & \text{subject to } x \in \mathbb{R}_*^n. \end{aligned}$$

ここで、 $\mathbb{R}_*^n = \mathbb{R}^n - \{0\}$ である。この問題の目的関数 $f(x) = x^T A x / x^T x$ はレイリー商と呼ばれる。目的関数 f の点 x における（各変数についての偏微分を並べたベクトルとして定義される）勾配 $\text{grad } f(x)$ およびヘッセ行列 $\text{Hess } f(x)$ を計算すると、それぞれ

$$\text{grad } f(x) = \frac{2}{x^T x} (A - f(x)I_n)x, \quad (2.1)$$

$$\text{Hess } f(x) = \frac{2}{x^T x} \left(I_n - \frac{2xx^T}{x^T x} \right) (A - f(x)I_n) \left(I_n - \frac{2xx^T}{x^T x} \right) \quad (2.2)$$

となる。したがって、この問題に対する点 x におけるニュートン方程式は、未知のベクトルを $d \in \mathbb{R}^n$ として、

$$\text{Hess } f(x)[d] = -\text{grad } f(x), \quad (2.3)$$

つまり、

$$\frac{2}{x^T x} \left(I_n - \frac{2xx^T}{x^T x} \right) (A - f(x)I) \left(I_n - \frac{2xx^T}{x^T x} \right) d = -\frac{2}{x^T x} (A - f(x)I_n)x \quad (2.4)$$

である。 x が f の停留点でないとき、この方程式の解は $d = x$ となる。つまり、原点と点 x を結ぶ直線上に f の停留点 x_* がない限り、ニュートン法による反復を何度行っても決して x_* には到達しない。したがって、問題 2.1 にニュートン法を適用しても、一般には最適解を得ることができない。

レイリー商関数は、二次形式 $x^T A x$ を x の標準ノルムの二乗 $x^T x$ で割ったものであるから、はじめから x のノルムは 1 であるという条件を付けておけば、レイリー商は $x^T A x$ に他ならない。よって、問題 2.1 を次のように書き換えることができる。

問題 2.2.

$$\begin{aligned} & \text{minimize } f(x) = x^T A x, \\ & \text{subject to } x \in \mathbb{R}^n, x^T x = 1. \end{aligned}$$

これは制約条件付きの最適化問題であり、たとえば拡張ラグランジュ法を用いて解くことができるが、収束は必ずしも速くない。問題 2.2 の制約条件を満たす点全体 $\{x \in \mathbb{R}^n \mid x^T x = 1\}$ は $n-1$ 次元球面 S^{n-1} であるから、さらに見方を変えて、この問題を球面 S^{n-1} 上の制約条件なしの最適化問題と見なすこともできる。

問題 2.3.

$$\begin{aligned} & \text{minimize } f(x) = x^T A x, \\ & \text{subject to } x \in S^{n-1}. \end{aligned}$$

こうすると、ユークリッド空間における最急降下法やニュートン法をリーマン多様体上に拡張しておけば、それを用いてこの問題を解くことができる。

2.2 レイリー商の重み付き和の最小化問題とシュティーフエル多様体上の最適化問題

次に、 $0 < \mu_1 < \dots < \mu_p$ を重みを表す定数として、レイリー商の p 個の重み付き和を最小化する問題を考える。ただし、 $p \leq n$ とし、また、 x_1, x_2, \dots, x_p は正規直交系をなすものとする。

問題 2.4.

$$\begin{aligned} & \text{minimize } \sum_{i=1}^p \mu_i x_i^T A x_i, \\ & \text{subject to } x_1, \dots, x_p \in \mathbb{R}_*^n, x_i^T x_j = \delta_{ij}. \end{aligned}$$

この問題の最適解 x_1^*, \dots, x_p^* は、それぞれ、行列 A の小さい方から p 個の固有値に属する p 個の固有ベクトル v_1, \dots, v_p である：

$$x_i^* = v_i, \quad i = 1, \dots, p. \quad (2.5)$$

もちろん、 A の小さい方から p 番目の固有値 λ_p と $p+1$ 番目の固有値 λ_{p+1} が等しいときは、最適解は一意には決まらない。

さて、行列 $N = \text{diag}(\mu_1, \dots, \mu_p)$ と、 x_1, \dots, x_p を並べてできる行列 $Y = (x_1, \dots, x_p)$ を導入すると、この問題の目的関数と制約条件はそれぞれ $F(Y) = \text{tr}(Y^T A Y N)$, $Y^T Y = I_p$ と書けるから、次のように行列を変数とする最適化問題であると見なすことができる。

問題 2.5.

$$\begin{aligned} & \text{minimize } F(Y) = \text{tr}(Y^T A Y N), \\ & \text{subject to } Y \in \mathbb{R}^{n \times p}, Y^T Y = I_p. \end{aligned}$$

この問題の実行可能領域 $\{Y \in \mathbb{R}^{n \times p} | Y^T Y = I_p\}$ には多様体の構造が入る。これをシュティーフエル多様体といい、 $\text{St}(p, n)$ と表す。すると、問題 2.5 は次のように書き換えられる。

問題 2.6.

$$\begin{aligned} & \text{minimize } F(Y) = \text{tr}(Y^T A Y N), \\ & \text{subject to } Y \in \text{St}(p, n). \end{aligned}$$

2.3 レイリー商の和の最小化問題とグラスマン多様体上の最適化問題

この節の最後に、次のレイリー商の p 個の和を最小化する問題を考える。

問題 2.7.

$$\begin{aligned} & \text{minimize } \sum_{i=1}^p x_i^T A x_i, \\ & \text{subject to } x_1, \dots, x_p \in \mathbb{R}^n, x_i^T x_j = \delta_{ij}. \end{aligned}$$

この問題の最小解 x_1^*, \dots, x_p^* は、行列 A の小さい方から p 個の固有値に属する p 個の固有ベクトル v_1, \dots, v_p と張る空間が一致する：

$$\text{span} \{x_1^*, \dots, x_p^*\} = \text{span} \{v_1, \dots, v_p\}. \quad (2.6)$$

この問題もまた、次のようにシュティーフエル多様体 $\text{St}(p, n)$ 上の問題に書き換えることができる。

問題 2.8.

$$\begin{aligned} & \text{minimize } F(Y) = \text{tr}(Y^T A Y), \\ & \text{subject to } Y \in \text{St}(p, n). \end{aligned}$$

ところで、目的関数 F は $O(p)$ 不変である。実際、 $Q \in O(p)$ に対して、

$$\begin{aligned} F(YQ) &= \text{tr}((YQ)^T A(YQ)) \\ &= \text{tr}(Q^T Y^T A Y Q) \\ &= \text{tr}(Y^T A Y Q Q^T) \\ &= \text{tr}(Y^T A Y) = F(Y). \end{aligned} \tag{2.7}$$

よって、探索領域を $\text{St}(p, n)/O(p)$ としても良い。 $\text{St}(p, n)/O(p)$ は \mathbb{R}^n の p 次元部分空間全体と同相であり、これにも多様体の構造が入る。この多様体をグラスマン多様体といい、 $\text{Grass}(p, n)$ と表す。

すると、問題 2.8 は次のように書くことができる。

問題 2.9.

$$\begin{aligned} &\text{minimize } F(Y) = \text{tr}(Y^T A Y), \\ &\text{subject to } [Y] \in \text{Grass}(p, n). \end{aligned}$$

つまり、問題 2.8 をグラスマン多様体上の制約条件なしの最適化問題に帰着することができた。

2.4 本稿でのグラスマン多様体の取り扱い

グラスマン多様体 $\text{Grass}(p, n)$ について、 $\text{Grass}(p, n) \simeq \text{St}(p, n)/O(p) \simeq \mathbb{R}_*^{n \times p}/GL(p)$ が知られており、グラスマン多様体をこれらの商多様体と見なして最適化アルゴリズムが導出されてきたが、本稿では

$$\begin{aligned} \text{Grass}(p, n) &\simeq \{X \in \mathbb{R}^{n \times n} \mid X^2 = X, X^T = X, \text{rank}(X) = p\} \\ &= \{X = Y Y^T \mid Y \in \text{St}(p, n)\} \end{aligned}$$

であることを用いて最適化アルゴリズムを導出する。つまり、グラスマン多様体上の各点を階数 p の直交射影行列と見なす。すると、問題 2.9 は次のように書くことができる。

問題 2.10.

$$\begin{aligned} &\text{minimize } F(X) = \text{tr}(A X), \\ &\text{subject to } X \in \text{Grass}(p, n). \end{aligned}$$

この問題の解法アルゴリズムを導出するための準備として、次の節ではリーマン多様体上の最適化アルゴリズムの一般論を説明する。

3 リーマン多様体上の最適化手法

3.1 ユークリッド空間における最適化手法をリーマン多様体上に拡張するための準備

まず最初に、次のユークリッド空間 \mathbb{R}^N における制約条件なしの最適化問題を考える。

問題 3.1.

$$\begin{aligned} & \text{minimize } f(x), \\ & \text{subject to } x \in \mathbb{R}^N. \end{aligned}$$

制約条件なしの問題に対しては、最急降下法やニュートン法などのアルゴリズムが知られているが、それらのアルゴリズムは基本的に次の枠組みを持つ。

Algorithm 3.1 \mathbb{R}^N 上の最適化アルゴリズム

- 1: 初期点 $x_0 \in \mathbb{R}^N$ を選ぶ.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: 探索方向 $\eta_k \in \mathbb{R}^N$ とステップサイズ $t_k > 0$ を求める.
 - 4: 次の点 x_{k+1} を $x_{k+1} := x_k + t_k \eta_k$ によって計算する.
 - 5: **end for**
-

探索方向（とステップサイズ）の決め方が各最適化アルゴリズムを特徴付ける。探索方向 $\eta_k \in \mathbb{R}^N$ は、たとえば最急降下法では x_k における f の逆勾配として、ニュートン法ではニュートン方程式の解として定める。

これらの概念をリーマン多様体上に拡張する。今から考える問題は、次のリーマン多様体 (M, g) 上の制約条件なしの最適化問題である。

問題 3.2.

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

まず、 k 反復目における探索方向 η_k は点 $x_k \in M$ における接ベクトルとして選ぶ。つまり、 $\eta_k \in T_{x_k} M$ となるようにする。

また、探索方向 $\eta_k \in T_{x_k} M$ とステップサイズ $t_k > 0$ を何らかの方法で求めたとして、ユークリッド空間上のアルゴリズムにおける更新の式

$$x_{k+1} := x_k + t_k \eta_k \tag{3.1}$$

は多様体上では一般には意味をなさない。実際、多様体はユークリッド空間に埋め込まれているとは限らないので加法が一般には定義されないし、仮にユークリッド空間 \mathbb{R}^N に埋め込まれていて $x_k + t_k \eta_k \in \mathbb{R}^N$ が定まるとしても、これは一般には多様体 M 上の点ではない。そこで、 $\gamma(0) = x_k$, $\dot{\gamma}(0) = \eta_k$ なる M 上の曲線 γ に沿って次の点 x_{k+1} を探索する。そのために、「探索をする上で妥当な」曲線を定める写像 $R: TM \rightarrow M$ が見つければ、 $R_x := R|_{T_x M}$ として、

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

なる更新式を得られる。

R に課すべき条件を考えよう。 $R_{x_k}: T_{x_k} M \rightarrow M$ によって M 上の曲線 $R_{x_k}(t\eta_k)$ を定義したときに、この曲線が x_k から η_k の方向に伸びるようにしたいから、

$$R_{x_k}(0) = x_k \quad (3.2)$$

かつ

$$\frac{d}{dt} R_{x_k}(t\eta_k)|_{t=0} = \eta_k \quad (3.3)$$

であるべきである。

$$\frac{d}{dt} R_{x_k}(t\eta_k)|_{t=0} = DR_{x_k}(0)[\eta_k] \quad (3.4)$$

であることと、これらの式が、点 x_k がいかなる点であっても、また探索方向 η_k がいかなるベクトルであっても成り立つべきであることに注意して、

$$R_x(0) = x, \quad DR_x(0)[\eta] = \eta, \quad x \in M, \eta \in T_x M \quad (3.5)$$

を得る。これらを満たす写像を、レトラクションと呼ぶことにする。

3.2 レトラクション

以上を踏まえて、レトラクション $R: TM \rightarrow M$ をあらためて次のように定義する。

定義 3.1. 写像 $R: TM \rightarrow M$ が次を満たすとき、 R をレトラクションという。 $R_x := R|_{T_x M}$ とする。

- $R_x(0_x) = x$, ここで、 0_x は $T_x M$ の零ベクトルである。
- $T_{0_x} T_x M \simeq T_x M$ なる同一視の下で、

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

図1は、 M がユークリッド空間 \mathbb{R}^N の部分多様体である場合のレトラクションの概念図である。

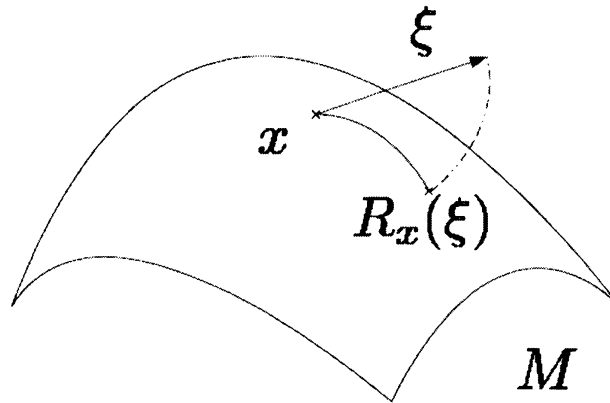


図 3.1: レトラクションの概念図

3.3 リーマン多様体上の最適化アルゴリズム

リーマン多様体 M 上のレトラクション R を用いて、 M 上の最適化アルゴリズムは次のように記述される：

Algorithm 3.2 多様体上の最適化アルゴリズム

初期点 $x_0 \in M$ を選ぶ.

for $k = 0, 1, 2, \dots$ **do**

探索方向 $\eta_k \in T_{x_k} M$ とステップサイズ $t_k > 0$ を求める.

次の点 x_{k+1} を, $x_{k+1} := R_{x_k}(t_k \eta_k)$ によって計算する.

end for

図 2 は、このアルゴリズムにおける、レトラクションにより定義される曲線に沿った探索を図示したものである。

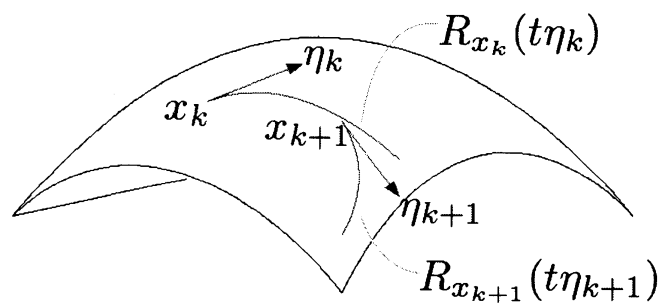


図 3.2: レトラクションにより定義される曲線上での探索

3.4 アルミホのルール

ステップサイズの決め方としては、アルミホの方法が有名であるが、これもリーマン多様体上に拡張することができる。

定義 3.2. f をリーマン多様体 M 上の目的関数とし、 R をレトラクションとする。スカラー $\bar{\alpha} > 0$, $\beta, \sigma \in (0, 1)$ は与えられた定数とする。

与えられた点 $x \in M$ と接ベクトル $\eta \in T_x M$ に対して、アルミホポイントを $\eta^A = \beta^m \bar{\alpha} \eta$ によって定義する。ここで、 m は

$$f(x) - f(R_x(\beta^m \bar{\alpha} \eta)) \geq -\sigma \langle \text{grad } f(x), \beta^m \bar{\alpha} \eta \rangle_x.$$

を満たす最小の非負の整数である。

また、 $t^A = \beta^m \bar{\alpha}$ をアルミホステップサイズという。

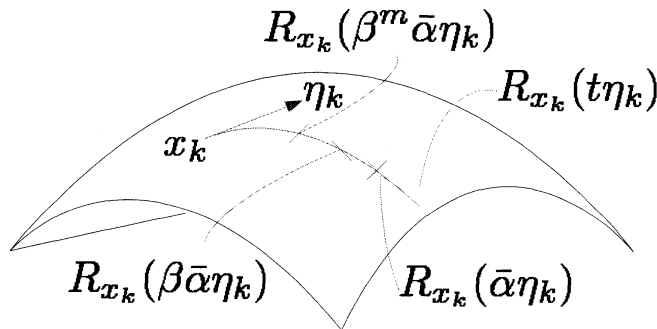


図 3.3: アルミホのルールによるステップサイズの決め方

3.5 最急降下法とニュートン法

リーマン多様体 (M, g) 上の最急降下法では、点 $x_k \in M$ における探索方向 $\eta_k \in T_{x_k} M$ を、

$$\eta_k = -\text{grad } f(x_k) \quad (3.6)$$

によって定める。ここで、 $\text{grad } f(x_k)$ は、点 x_k における、リーマン計量 g に関する f の勾配である。

一方ニュートン法では、探索方向 η_k を、ニュートン方程式

$$\text{Hess } f(x_k)[\eta_k] = -\text{grad } f(x_k) \quad (3.7)$$

の解として定める。なお、本稿ではヘシアン $\text{Hess } f$ はレヴィ・チビタ接続により定義されるものを用いる。

4 グラスマン多様体とその上のレイリー商の幾何

4.1 レトラクション

グラスマン多様体上のレトラクションを2通り述べる. 一つは指数写像によるものであり, もう一つはQR分解を用いて定義されるものである. これらをそれぞれ指数レトラクション, QRレトラクションと呼ぶことにする.

指数レトラクション

指数写像を求めるためには測地線方程式を解けば良いが, 測地線方程式はグラスマン多様体上の自由粒子を考え, 最小作用の原理を適用することにより導出することができる. その過程は省略するが, 測地線方程式は

$$\ddot{X} + 2\dot{X}^2 - 4\dot{X}X\dot{X} = 0. \quad (4.1)$$

となり, その初期条件 $X(0) = X, \dot{X}(0) = \xi$ の下での解 $X(t) =: \text{Exp}_X(t\xi)$ は,

$$\text{Exp}_X(t\xi) = \frac{1}{2} \begin{pmatrix} YV & U \end{pmatrix} \begin{pmatrix} I + \cos(2\Sigma t) & \sin(2\Sigma t) \\ \sin(2\Sigma t) & I - \cos(2\Sigma t) \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix} \quad (4.2)$$

となる. ここで, $U \in \text{St}(p, n)$, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_p)$, $V \in O(p)$ は ξY の特異値分解である;

$$\xi Y = U \Sigma V^T, \quad \sigma_1 \geq \dots \geq \sigma_p \geq 0. \quad (4.3)$$

ただし, $X = Y Y^T$, $Y \in \text{St}(p, n)$ である ($X \in \text{Grass}(p, n)$ に対して, 必ずこのような Y が存在する).

以上から, 指数写像は

$$\text{Exp}_X(\xi) = \frac{1}{2} \begin{pmatrix} YV & U \end{pmatrix} \begin{pmatrix} I + \cos(2\Sigma) & \sin(2\Sigma) \\ \sin(2\Sigma) & I - \cos(2\Sigma) \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix} \quad (4.4)$$

と書ける. 指数写像を用いて,

$$R := \text{Exp} \quad (4.5)$$

によって $R: T\text{Grass}(p, n) \rightarrow \mathbb{R}$ を定義すると, これはレトラクションである. この R を指数レトラクションと呼ぶことにする.

QRレトラクション

$n \times n$ 行列 B に対して, $\text{qf}(B)$ を, 行列 B のQR分解の Q 成分とする. つまり, B が

$$B = QR, \quad Q \in O(n), R \in S_{\text{upp}}^+(n),$$

のように分解されるとき, $\text{qf}(B) = Q$. ここで, $S_{\text{upp}}^+(n)$ は対角成分が正の上三角行列全体からなる集合である.

$X \in \text{Grass}(p, n)$ に対して $Y \in \text{St}(p, n)$ を $X = YY^T$ を満たすものとして選ぶ. QR 分解を用いて

$$R_X(\xi) = \text{qf}((I + \xi)Y) \text{qf}((I + \xi)Y)^T$$

と定義すると, これも $\text{Grass}(p, n)$ 上のレトラクションであることが確かめられる. この R を QR レトラクションと呼ぶことにする.

4.2 レイリー商の勾配とヘシアン

以後, 次のグラスマン多様体上のレイリー商最小化問題について考える.

問題 4.1.

$$\begin{aligned} & \text{minimize } F(X) = \frac{1}{2} \text{tr}(AX), \\ & \text{subject to } X \in \text{Grass}(p, n). \end{aligned}$$

$F(X) = \text{tr}(AX)$ について, 勾配とヘシアンを定義に従って計算すると, 次のようになる. まず $X \in \text{Grass}(p, n)$ における勾配 $\text{grad } F(X)$ は,

$$\text{grad } F(X) = \text{sym}(AX) - XAX \quad (4.6)$$

である. ここで, $\text{sym}(D) = (D + D^T)/2$ である. また, $X \in \text{Grass}(p, n)$ におけるヘシアン $\text{Hess } F(X)$ は, $\xi \in T_X \text{Grass}(p, n)$ に対して

$$\text{Hess } F(X)[\xi] = 2 \text{sym}((X\xi A - XA\xi)(I - X)) \quad (4.7)$$

のように作用する.

4.3 最急降下法とニュートン法・ハイブリッド法

こうして問題 4.1 に対する解法アルゴリズムを書き下すことができる. ステップサイズはアルミホステップサイズを採用すると決めておけば, 探索方向 η_k の計算方法さえ決めてしまえば, アルゴリズム 3.2 を実装することができる. 特に, ニュートン法においては, ニュートン方程式が

$$\xi_k B_k + B_k \xi_k = C_k \quad (4.8)$$

となる. ここで, $B_k := A - AX_k - X_k A$, $C_k := 2(-\text{sym}(AX_k) + X_k AX_k)$ は, 方程式 (4.8) を解く際には定数行列である. 方程式 (4.8) はリヤプノフ方程式と呼ばれるものであり, 解法がよく研究されている [3].

リーマン多様体 M 上の最急降下法とニュートン法について、次のことが知られている。まず、 M がコンパクトであれば、最急降下法によって生成される列 $\{x_k\}_{k=0,1,\dots}$ は、目的関数 f の臨界点に線形収束する。一方、目的関数 f の臨界点 $x_* \in M$ において $\text{Hess } f(x_*)^{-1}$ が存在するとき、 x_* の十分近くを初期点にとると、ニュートン法によって生成される点列 $\{x_k\}$ は x_* に2次収束する。また、証明は割愛するが、グラスマン多様体上のレイリー商関数に関しては、大域的最大点と大域的最小点以外の臨界点はすべて鞍点であることが示せた。したがって、この場合は最急降下法による点列は大域的最適解に大域的に収束する。そこで、最急降下法とニュートン法を組み合わせたハイブリッド法の有効性が期待されるが、最急降下法からニュートン法に切り替えるタイミングを決める基準を定めるのが困難であり、それは今後の課題である。

5 終わりに

以上述べてきた多様体上の最適化アルゴリズムによって、グラスマン多様体上のレイリー商最小化問題を、ユークリッド空間上の制約条件付きの最適化問題と見なした場合の解法より効率良く解くことができるが、この問題を固有値問題であると見なした場合にはより良い既存のアルゴリズムが多くあり、提案手法は固有値問題の解法としては未だ実用的ではない。より速いアルゴリズムの導出も今後の課題の一つである。

なお、多様体上の最適化手法の詳細については [1, 2, 5] を、QR 分解や特異値分解といった行列の数値的な話題については [4] を、多様体などの微分幾何学の詳細については [6] を参照されたい。

参考文献

- [1] P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, Princeton University Press, Princeton, NJ, 2008.
- [2] A. Edelman, T.A. Arias, S.T. Smith, The geometry of algorithms with orthogonality constraints, SIAM J. Matrix Anal. Appl. 20 (2) (1998) 303-353.
- [3] Z. Gajic, M.T.J. Qureshi, Lyapunov Matrix Equation in System Stability and Control, Academic Press, 1995.
- [4] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 1996.
- [5] U. Helmke, J.B. Moore, Optimization and Dynamical Systems, in: Communications and Control Engineering, Springer, London, 1994.
- [6] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, Interscience Publishers, New York-London, 1963. Volumes 1 and 2.