

Optimization Algorithm on Riemannian Manifolds

小泉孝弥

立命館大学大学院 修士 2 年

April 15, 2022

Contents

1 Introduction

2 Riemannian Manifold and Examples

- Definition and Gradient
- Examples

3 Extension to Riemannian Manifold

■ Retraction

■ Riemannian Gradient Decent Method

4 Numerical Experiments

- Minimization of quadratic form

Introduction

本発表では制約空間 M がリーマン多様体となっている非線形連続最小化問題を扱う. すなわち, 関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が与えられたとき,

$$\begin{array}{|l} \text{minimize} \quad f(x) \\ \text{subject to } x \in M \end{array}$$

制約空間 M がリーマン多様体となっている場合, f を M に制限することで, 制約なしの問題と考えることができる. 今回は, 勾配降下法をリーマン多様体上に拡張し, 数値計算を行なった結果を報告する.

Contents

1 Introduction

2 Riemannian Manifold and Examples

- Definition and Gradient
- Examples

3 Extension to Riemannian Manifold

■ Retraction

■ Riemannian Gradient Decent Method

4 Numerical Experiments

- Minimization of quadratic form

Riemannian Manifold

Definition (Riemannian Manifold)

M を可微分多様体とする. 各点 $x \in M$ の接空間 $T_x M$ に内積 $g_x : T_x M \times T_x M \rightarrow \mathbb{R}$ が定まっているとき, $g = (g_x)_{x \in M}$ をリーマン計量と呼び, 組 (M, g) をリーマン多様体 (Riemannian Manifold) という.

Definition (Gradient)

(M, g) をリーマン多様体とし, $f : M \rightarrow \mathbb{R}$ を可微分関数とする. $x \in M$ について f の x での勾配 $\text{grad} f(x) \in T_x M$ を以下で定める.

$$\forall \xi \in T_x M, g(\text{grad} f(x), \xi) = Df(x)[\xi]$$

Sphere

Example (Sphere)

自然数 $n \geq 2$ に対して, 球面 $S^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ は可微分多様体となる. また, $x \in S^{n-1}$ での接空間 $T_x S^{n-1}$ は

$$T_x S^{n-1} = \{z \in \mathbb{R}^n \mid x^\top z = 0\}$$

となる. ここで, $g_x : T_x S^{n-1} \times T_x S^{n-1} \rightarrow \mathbb{R}$ を

$$g_x(\xi, \eta) = \xi^\top \eta$$

と定めれば, g_x は $T_x S^{n-1}$ の内積となるので, S^{n-1} はリーマン多様体である.

Stiefel Manifold

Example (Stiefel Manifold)

$X = [x_1 x_2 \dots x_p] \in \mathbb{R}^{n \times p} (n \geq p)$ で, $\{x_i\}_{i=1}^p \subset \mathbb{R}^n$ が正規直交系であるような $n \times p$ 行列全体は可微分多様体となる. この多様体をシュティーフエル多様体 (*Stiefel Manifold*) といい, $\text{St}(p, n)$ と表す. すなわち

$$\text{St}(p, n) = \{X \in \mathbb{R}^{n \times p} \mid X^\top X = \mathbb{I}_p\}.$$

である.

$X \in \text{St}(p, n)$ での接空間 $T_X \text{St}(p, n)$ は,

$$\begin{aligned} T_X \text{St}(p, n) &= \left\{ Z \in \mathbb{R}^{n \times p} \mid X^T Z + Z^T X = 0 \right\} \\ &= \left\{ X\Omega + X_\perp K \mid \Omega \in \text{Skew}(p), K \in \mathbb{R}^{(n-p) \times p} \right\} \end{aligned}$$

となる. ここで $\text{Skew}(p) = \{\Omega \in \mathbb{R}^{p \times p} \mid \Omega^T = -\Omega\}$ であり, X_\perp は $\text{span}(X)^\perp = \text{span}(X_\perp)$ を満たす $n \times (n-p)$ 行列である.
 $g_X : T_X \text{St}(p, n) \times T_X \text{St}(p, n) \rightarrow \mathbb{R}$ を,

$$g_X(\xi, \eta) = \text{tr}(\xi^T \eta)$$

と定めれば, g_X は内積となるので, $\text{St}(p, n)$ はリーマン多様体である.

Contents

1 Introduction

2 Riemannian Manifold and Examples

- Definition and Gradient
- Examples

3 Extension to Riemannian Manifold

■ Retraction

■ Riemannian Gradient Decent Method

4 Numerical Experiments

- Minimization of quadratic form

Gradient Decent

Algorithm 1 Gradient Decent(GD)

Require: f : differentiable function on \mathbb{R}^n

Require: $0 < t < 1$:

$x \leftarrow x_0 \in \mathbb{R}^n$

while x not converged **do**

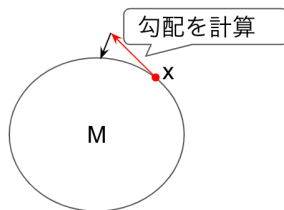
$x \leftarrow x - t \operatorname{grad}_{\mathbb{R}^n} f(x) \in \mathbb{R}^n$

end while

return x

Motivation

勾配降下法を制約空間がユークリッド空間に埋め込まれたリーマン部分多様体に拡張する。その際に、勾配の部分リーマン多様体上の勾配に変更するだけでは、リーマン多様体からはみ出してしまうため、制約条件を満たすことが出来ない。そこで、はみ出ってしまった点をリーマン多様体に戻す写像を考えることで、制約を満たす点列 $\{x_k\}$ を得ることができる。



Gradient Decent

Algorithm 2 Gradient Decent(GD)

Require: f : differentiable function on \mathbb{R}^n

Require: $0 < t < 1$:

$x \leftarrow x_0 \in M$

while x not converged **do**

$x \leftarrow x - t \operatorname{grad} f(x) \notin M$

end while

return x

Retraction

リーマン多様体 M に対して, 接束を $TM = \bigcup_{x \in M} T_x M$ と表す.

Definition (Retraction)

M をリーマン多様体, $R: TM \rightarrow M$ を C^∞ 写像とする. 任意の $x \in M$ に対して, R の $T_x M$ への制限 $R_x: T_x M \rightarrow M$ が以下の 2 つの条件を満たすとき, R を M 上のレトラクション (Retraction) と呼ぶ.

- 1 $R_x(0_x) = x$
- 2 $DR_x(0_x) = id_{T_x M}$

Riemannian Gradient Decent

M をリーマン多様体, $R : TM \rightarrow M$ をレトラクションとする.
リーマン多様体上の勾配降下法 (RGD) を以下のように設計する.

Algorithm 3 Riemannian Gradient Decent(RGD)

Require: f : differentiable function on M

Require: $0 < t < 1$:

$x \leftarrow x_0 \in M$

while x not converged **do**

$x \leftarrow R_x(-t \operatorname{grad} f(x))$

end while

return x

Contents

- 1 Introduction
- 2 Riemannian Manifold and Examples
 - Definition and Gradient
 - Examples
- 3 Extension to Riemannian Manifold
- Retraction
- Riemannian Gradient Decent Method
- 4 Numerical Experiments
 - Minimization of quadratic form

二次形式の最小化

二次形式の最小化問題

$A \in \mathbb{R}^{n \times n}$ を対称行列とする.

$$\begin{array}{ll} \text{minimize} & q_A(x) = x^\top A x \\ \text{subject to} & x \in S^{n-1} \end{array}$$

なお, $\min_{x \in S^{n-1}} q_A(x)$ は A の最小固有値, $\arg \min_{x \in S^{n-1}} q_A(x)$ は最小固有値に対する固有ベクトルとなる.

Result

レトラクション R を以下で定義する.

$$R_x(\xi) = \frac{x + \xi}{\|x + \xi\|}$$

Figure: $n = 100$

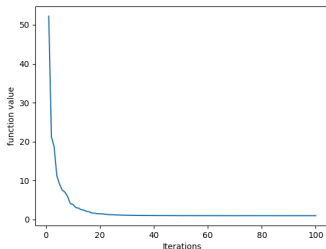
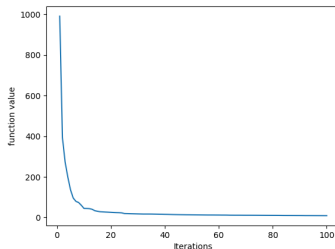


Figure: $n = 2000$



Brocket cost の最小化

Brocket cost minimization

$C \in \mathbb{R}^{n \times n}$ を対称行列, $N = \text{diag}(\mu_1, \mu_2, \dots, \mu_p)$ とする (ただし, $0 < \mu_1 < \mu_2 < \dots < \mu_p$).

$$\begin{array}{|l} \text{minimize} \\ \text{subject to} \end{array} \quad \begin{array}{l} F(X) = -\text{tr}(X^\top C X N) \\ X \in \text{St}(p, n) \end{array}$$

いわゆる主成分分析 (PCA). $p = 1$ の時は先程の問題とおなじになる.

Result

レトラクション R_X を以下で定義する.

$$R_X(\xi) := \text{Q factor of QR decomposition of } (X + \xi)$$

References

- [1] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE,
Optimization Algorithms on Matrix Manifolds, Princeton
University Press, Princeton, NJ, 2008.