1 Introduction

1.1 Abstract

最初に論文に記載されている内容を説明する. 数学的に示していることをざっくりとした (数学的には不正確な) 日本語で説明すると以下の二点である.

- 多様体仮説がある多様体に対して, 成り立つと仮定する. この時, 実際に多様体仮説が成り立つ多様体 M を数学的に構成した.
- 上の多様体 M を具体的に求めるアルゴリズムを構成した.

(この論文) では, 入力となるデータに関する性質を記述できていないので, 多様体仮説が成り立つことは確認できない. だが, 多様体仮説が成り立つかはわからないが, 多様体仮説が成り立つならば, 求めたい多様体 (といっていいもの) が定義でき, さらにそれを実際に求めるアルゴリズムを構成できたので, 実際に実験してみればよいと主張している. もし数学的に多様体仮説を示すのであれば, 入力データの偏りに相当する仮定が必要になる. また, 多様体仮説を聞いて, 多様体の形に基づいた, クラスタリングが定義できるのではないかと予想していたが, この論文では特に記述はなかった. (例えば, 連結成分等でうまく定義できて欲しいものだが...)

結果を具体的に記述するため, 最低限の数学的な用語を定義する.

1.2 Notation and Definition

Definition 1.1. 有限次元とは限らない $\mathcal{H}$ を $\mathbb{R}$-ベクトル空間が以下を満たす時, ヒルベルト空間という.

- 内積が定義できる. すなわち, 以下を満たす写像 $< \cdot, \cdot >: \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ が存在する.
    1. 任意の $x_1, x_2, y \in \mathcal{H}$ に対し,$<x_1 + x_2, y> = <x_1, y> + <x_2, y>$
    2. 任意の $x, y \in \mathcal{H}, a \in \mathbb{R}$ に対し,$< ax, y >= a < x, y >$
    3. 任意の $x, y \in \mathcal{H}$ に対し,$< x, y >=< y, x >$
    4. 任意の $x \in \mathcal{H}$ に対し,$<x, x> \geq 0$ であり,$<x, x> = 0$ と $x = 0$ は同値である.
- 内積が定める距離について完備である.

Remark. 上で定めた内積のことを正定値対象双線型形式という.

Remark. $x, y \in \mathcal{H}$ に対し,$d(x, y) := \sqrt{<x - y, x - y>}$ とすると距離の公理を満たす. 距離の公理とは以下の 3 つのこと.

1. $d(x,y) = d(y,x)$

2. $d(x,y) = 0$ は $x = y$ と同値

3. $d(x,y) + d(y,z) \geq d(x,z)$

Remark. 距離空間 X が完備とは, 任意のコーシー列 $\{x_n\}_{n \in \mathbb{Z}_{\geq 0}}$ が, ある $x \in X$ に対する収束列となること.

Example 1.2. $\mathbb{R}^n$ は標準内積により, ヒルベルト空間になる.

この論文で使われる"多様体"の定義をする.

Definition 1.3 (reach)**.** M を $\mathcal{H}$ の部分集合とし.$x \in \mathcal{H}$ に対し,$d(x,M)$ を $\inf_{y \in M} d(x,y)$ で定める. 実数 τ が M の ***reach*** とは,$d(x,M) < \tau$ となる任意の x に対し, $y \in M$ が存在し,y と異なる任意の $z \in M$ に対し,$d(x,z) > d(x,M) = d(x,y)$ となること.

Example 1.4. 半径 r の円周 $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$ の *reach* は r になる.

Remark. *reach* の定義は初めて見た. 恐らく多様体をクラスタリングするために, 定義されているのだと思う.

Definition 1.5. 位相空間 X の部分集合 Y が**稠密**とは Y の閉包が X に一致することをいう. 位相空間 M が稠密な加算部分集合を持つ時, **可分**という.

Example 1.6. 実数体 $\mathbb{R}$ に通常のユークリッド位相を定めた時, 有理数体 $\mathbb{Q}$ は稠密になる. また $\mathbb{Q}$ は加算集合のため,$\mathbb{R}$ は可分である.

Definition 1.7 (Tangent Space)**.** H を可分ヒルベルト空間とする. 閉集合 $A \subset \mathcal{H}$ と $a \in A$ に対し,$Tan^0(a,A)$ を以下で定める.

$$\{v \in \mathcal{H} \mid \text{任意の}\epsilon\text{に対し, ある}\, b \in A\, \text{が存在し,}\, 0 < |b-a| < \epsilon \text{に対し} \, |\frac{v}{|v|} - \frac{b-a}{|b-a|} < \epsilon.\}$$

$Tan(a,A)$ を $\{x \in H | x - a \in Tan^0(a,A)\}$ で定め,a での ***Tangent Space*** という.

Propostion 1.8. A を $\mathbb{R}^n$ の閉部分集合とする. この時

$$reach(A)^{-1} = \sup\{2|b-a|^{-2} d(b, Tan(a,A)) | a.b \in A\}$$

Definition 1.9. $V \subset \mathcal{H}$ が d 次元**アフィン空間**とは, ある $a \in \mathcal{H}$ が存在し,$\{x \in \mathcal{H} \mid x + a \in V\}$ が d 次元のベクトル空間になること.

Definition 1.10 (C^r-submfd)**.** ヒルベルト空間 $\mathcal{H}$ の閉集合 M は以下を満たす時, d 次元 C^r **級部分多様体**という.

- 任意の $p \in M$ に対し,$x \in U$ となる $\mathcal{H}$ の開部分集合と C^r 級写像 $\phi : U \to \mathcal{H}$ が存在する.
- $\phi|_{U \cap M}$ は終域を像に制限すると微分同相写像である
- d 次元アフィン空間の族 $\{V_i\}_{i \in I}$ に対し,$\phi(U \cap M) = \cap_{i \in I} V_i \cap \phi(U)$ が成り立つ.

$B_{\mathcal{H}} := \{x \in \mathcal{H} \mid d(x,x) \leq 1\}$ とする. $\mathcal{G}(d,V,\tau)$ を体積 V 以下であり,*reach* が τ 以上となる $B_{\mathcal{H}}$ に含まれる d 次元 *(*境界のない*)*C^r 級多様体全体のなす集合とする.

Remark. 多様体は普通の数学書では上記のようには定義されない. $\mathcal{H}$ を固定して考えたいがために, 上記のように定義したのだと思われる.

Remark. この論文では, 後でもう一つ多様体と呼ぶものが出てくる.C^r 級部分多様体が必ずしも後で出てくる多様体とは限らないことに注意せよ.

Definition 1.11. Ω を集合とする. F を Ω 上の σ-加法族, μ_P を F 上の非負測度で,$\mu_P(\Omega) = 1$ とする. この時,(Ω, F, μ_P) を確率空間という.

Definition 1.12. $X : \Omega \to [-\infty, \infty]$ が確率変数とは任意の $a \in [-\infty, \infty]$ で,$\{x \in \Omega \mid X(x) \leq a\} \in F$ となること.

Definition 1.13. X を確率変数とした時,$P(x) = \mu_P(X < x)$ と書ける時,P を確率分布関数という.

1.3 Manifold Hypothesis

この論文で考える多様体仮説とそのために必要な用語を定義する. 可分なヒルベルト空間を $\mathcal{H}$, その単位球を $B_{\mathcal{H}}$, $B_{\mathcal{H}}$ 上の確率分布関数を P とする.

$$\mathcal{L}(M,P) := \int_{x \in B_{\mathcal{H}}} d(x,M)^2 \frac{dP(x)}{dx} dx.$$

とする. 確率分布関数 P による i.i.d(独立試行) が行われた時, 以下が成り立つことを多様体仮説と呼ぶ.

HYPOTHESIS 1. ある $M \in \mathcal{G}(d, CV, \tau/C)$ が存在し,$\mathcal{L}(M,P) \leq C\epsilon$ となる.

Remark. 論文中でこれを多様体仮説と直接呼んではいないが, 内容から判断して多様体仮説と呼ぶことにした.

Remark. この時点でもわかるように,d, C, V, τ としてどのような値を取るべきかは不明.

1.4 Main Theorem 1

この論文の 1 つめの主定理を述べる.

Main Thoerem 1. ある $r > 0$ に対し,

$$U_{\mathcal{G}(1/r)} := CV(\frac{1}{\tau^d} + \frac{1}{(\tau r)^{d/2}})$$

とする. また,

$$s_{\mathcal{G}}(\epsilon, \delta) := C(\frac{U_{\mathcal{G}}(1/\epsilon)}{\epsilon^2}(\log^4(\frac{U_{\mathcal{G}}(1/\epsilon)}{\epsilon})) + \frac{1}{\epsilon^2}\log\frac{1}{\delta})$$

とする. $s \geq s_{\mathcal{G}}(\epsilon, \delta)$ とし,$x = \{x_1, \ldots, x_s\}$ を確率分布 P による独立試行により, 得られた集合とする. P_X を X 上の一様確率分布関数とする. $M_{erm}(x) \in \mathcal{G}(d, V, \tau)$ を

$$L(M_{erm}(x), P_X) - \inf_{M \in \mathcal{G}(d,V,\tau)} L(M, P_X) < \frac{\epsilon}{2}$$

を満たす $\mathcal{G}(d, V, \tau)$ で, 以下を最小にするものとする.

$$\sum_{i=1}^{s} d(x_i, M)^2$$

この時,

$$\mathbb{P}[L(M_{erm}(x), P) - \inf_{M \in \mathcal{G}(d,V,\tau)} L(M, P) < \epsilon] > 1 - \delta$$

となる.

Remark. s 次元の直積測度で考えた時, 上の条件を満たす x 全体の集合をその測度に代入すると $1 - \delta$ より大きいという意味.

Remark. この定理が何を示しているかというと,$L(M, P)$ の極小値に十分近い値を取る, 多様体 $M_{erm}(x)$ を構成できることが言えた.$M_{erm}(x)$ は最適化問題の解として得られるので, 極小値よりもかなり具体的になっている. 特に,$M_{erm}(x)$ が確率を計算しなてくも求められるのがよい. 基本的に P_X は未知であることを前提に問題と解くので,P_X に依存しないのは強い結果である.

Remark. 仮定からわかるように,$\epsilon.\delta$ の値にも強く依存する. 私はまだ, これらとして何を取るべきかはわかっていない. 計算してみて, どのぐらい大きい s が必要か推測 *(*計算*)* したい.

2 Preparation for Main Theorem 1

主定理 1 を示すための準備を行う. 多様体を考察することで以下を示す.

Claim 1. *$M \in \mathcal{G}(d, V, \tau)$ とし,Π_x を $\mathcal{H} \to \mathrm{Tan}(x, M)$ への射影とする．十分大きい (control された?) 定数 C に対し,*

$$U := \{y \mid |y - \Pi_x y| \leq \tau/C\} \cap \{y \mid |x - \Pi_x y| \leq \tau/C\}$$

とする．この時,$C^{1,1}$ 級関数 $F_{x,U} : \Pi_x(U) \to \Pi_x^{-1}(\Pi_x(0))$ で以下を満たすものが存在する．

1. *$F_{x,U}$ の Lipschitz constant of the gradient が C 以下である．*
2. *$M \cap U = \{y + F_{x,U}(y) \mid y \in \Pi_x(U)\}$*

2.1 Proof of Claim 1

$\mathcal{H}$ を (無限次元でもよい) ヒルベルト空間とする．D-plane を $\mathcal{H}$ の D 次元部分ベクトル空間とする．DPL で D-plane 全体のなす空間を表す．$\Pi, \Pi' \in DPL$ に対し,$T : \mathcal{H} \to \mathcal{H}$ を直交変換 (内積が不変な変換) であって,$T(\Pi) = \Pi'$ とを満たすとする.T 全体のなす集合を $A_{\Pi,\Pi'}$ とする．この時,

$$\mathrm{dist}(\Pi, \Pi') := \mathrm{inf}_{T \in A_{\Pi,\Pi'}} ||T - I||$$

で定める．$(DPL, dist)$ は距離空間になる．

D-palne Π に対し,

$$\Psi : B_\Pi(0, r) \to \Pi^\perp$$

を $C^{1,1}$ 級関数 (定義不明) で $\Psi(0) = 0$ とする．この時,a patch of radius r over Π centered at 0 とは,

$$\Gamma = \{x + \Psi(x) \mid x \in B_\Pi(0, r)\}$$

をさす. さらに

$$||\Gamma||_{C^{1,1}(B_\Pi(0,r))} := \mathrm{sup}_{x \neq y \in B_\Pi(0,r)} \frac{\nabla\Psi(x) - \nabla\Psi(y)}{||x - y||}$$

とする. ここで $\nabla\Psi(x) : \Pi \to \Pi^\perp$ は接ベクトル空間の間の射 $T_x\Pi \to T_x\Pi^\perp$ のこと (としか考えられない)．ただし,$\Pi \sim T_x\Pi, \Pi^\perp \sim T_x\Pi^\perp$ を使って同一視している．実質 x での gradient.

もし $\nabla\Psi(0) = 0$(0 写像) が成り立っている場合,Γ を a patch of radius r tangent to Π at its center 0 という．

Lemma 2.1. *Γ_1 を Π_1 上の半径 r_1 で中心 z_1 で Π_1 に接している patch とする．$z_2 \in \Gamma_1$ で $||z_2 - z_1|| < c_0 r_1$ を満たすとする．*

$$||\Gamma_1||_{C^{1,1}(B_\Pi(z_1,r_1))} \leq \frac{c_0}{r_1}$$

とする.$\Pi_2 \in DPL$ で $\mathrm{dist}(\Pi_2, \Pi_1) < c_0$ とする．この時,Π_2 上の半径 $c_1 r_1$ で中心 z_2 の *patch* Γ_2 で以下を満たすものが存在する．

$$||\Gamma_2||_{C^{1,1}(B_\Pi(0,c_1 r_1))} \geq \frac{200c_0}{r_1}$$

と

$$\Gamma_2 \cap B_{\mathcal{H}}(z_2, \frac{c_1 r_1}{2}) = \Gamma_1 \cap B_{\mathcal{H}}(z_2, \frac{c_1 r_1}{2})$$

Proof. わからん$\cdots$. 恐らく陰関数定理を使いたいのだろうが, 条件を満たす Γ の一意性が言えないような$\cdots$ $\square$

Definition 2.2. *$M \subset \mathcal{H}$ が"compact imbedded D-manifold" (for short, just a "manifold") とは, 以下が成り立つことである.*

- *M が compact*
- *任意の $z \in M$ に対し, ある $T_z M \in DPL$ と,$T_z M$ 上の半径 r_1, 中心 z,z で $T_z(M)$ に接する patchΓ が存在し，$M \cap B_{\mathcal{H}}(z, r_2) = \Gamma \cap B_{\mathcal{H}}(z, r_2)$ となる.*

We say that M has infinitesimal reach ρ if for every $\rho' < \rho$, there is a choice of $r_1 > r_2 > 0$ such that for every $z \in M$ there is a patch Γ over $T_z(M)$ of radius r_1, centered at z and tangent to $T_z(M)$ at z which has $C^{1,1}$-norm at most $1/\rho'$

Lemma 2.3 (Growing Patch)**.** *Let M be a manifold and let r_1, r_2 be as in the definition of a manifold. Suppose M has infinitesimal reach ≥ 1. Let $\Gamma \subset M$ be a patch of radius r centered at 0, over $T_0 M$. Suppose r is less than a small enough constant $\hat{c}$ determined by D. Then there exists a patch Γ^+ of radius $r + cr_2$ over $T_0 M$, centered at 0 such that $\Gamma \subset \Gamma^+ \subset M$*

Proof. $\mathcal{H}$ はヒルベルト空間なので,$\mathcal{H} = T_0 M \oplus T_0 M^\perp$ に直和分解する．M が D 次元なので,$T_0 M \simeq \mathbb{R}^D$ となる．patch Γ を $\Gamma = \{(x, \Psi(x) \mid x \in B_{\mathbb{R}^D}(0, r)\}$ と書く, ここで,$C^{1,1}$ 写像 $\Psi : B_{\mathbb{R}^D}(0, r) \to \mathcal{H}^\perp$ は $\Psi(0) = 0, \nabla\Psi(0) = 0$ となり,$||\Psi||_{C^{1,1}(B_{\mathbb{R}^D}(0,r))} \leq C_0$ となる．r を十分小さくとり,$y \in B_{\mathbb{R}^D(0,r)}$ に対し, $|\nabla\Psi(y)| \leq C_0||y|| \leq C_0$ とできる. この時, 上の lemma より (本当はきちんと条件を確認したい. $\Pi_1 = \Pi_2$ でとっているのと平行移動していること, そもそもこの取り方してるならもっと弱い補題でいけるのでは?)

$$\Psi_y : B_{\mathbb{R}^D}(y, c' r_2) \to \mathcal{H}^\perp$$

で $||\nabla\Psi_y(z)||, ||\Psi_y||_{C^{1,1}}$ が有界で,

$$M \cap B_{\mathcal{H}}((y, \Psi(y)), c'' r_2) = \{(z, \Psi_y(z)) \mid z \in B_{\mathbb{R}^D}(y, c' r_2)\} \cap B_{\mathcal{H}}((y, \Psi(y)), c'' r_2)$$

$\square$

3 A bound on the size of an ϵ-net

この章では、Claim 1 と経験過程の定理を使い、実際に主定理 1 を示す。

Definition 3.1. *Let* (X, d) *be a metric space, and* $r > 0$*. We call that* $Y \subset X$ *is an* r*-net of* X *if for each* $x \in X$*, there is a* $y \in Y$ *such that* $d(x, y) < r$*.*

Corllary 3.2. *Let*

$$U_{\mathcal{G}} : \mathbb{R}^+ \to \mathbb{R}$$

be given by

$$U_{\mathcal{G}}(1/r) = CV(\frac{1}{\tau^d} + \frac{1}{(\tau d)^{d/2}}).$$

(なぜ、逆数で与える。普通に違和感。) Let $M \in \mathcal{G}$*, and* M *be equipped with the metric* $d_{\mathcal{H}}$ *of the* $\mathcal{H}$*. Then, for any* $r > 0$*, there is an* $\sqrt{\tau r}$*-net of* M *consisting of no more than* $U_{\mathcal{G}}(1/r)$ *points.*