Manifold Denosing

본 문서는 [1] 논문을 간략히 요약한 것이다.

Introduction

The noise model and problem statement

데이터는 m dimensional abstract manifold M 위에 존재한다고 가정한다, 이 데이터는 smooth regular embedding $i : M \rightarrow \mathbb{R}^d$ 으로 Feature space $\mathbb{R}^d$ 에 embedding 된다. 즉,

$$i(M) \subset M \tag{1}$$

이때 data Generating process $X \in \mathbb{R}^d$ 는 다음과 같이 정의된다.

$$X = i(\Theta) + \epsilon \tag{2}$$

where $\theta \sim P_M$ and $\epsilon \sim N(0, \sigma)$. Probability measure P_M은 M 의에서 해당되는 Volume dV의 비율로 정의된다, 그러므로, $P_X(x)$는 P_M 으로 부터 다음과 같이 정의된다.

$$P_X(x) = (2\pi\sigma^2)^{-\frac{d}{2}} \int_M e^{-\frac{\|x-i(\theta)\|^2}{2\sigma^2}} p(\theta) dV(\theta) \tag{1}$$

- 이때 Volume dV는, Local coordinates 가 $\theta_1, \cdots, \theta_m$ 이면 다음과 같다.

$$dV = \sqrt{\det g} d\theta_1, \cdots d\theta_m \tag{3}$$

이때, $\det g$는 **metric tensor** g의 determinent 이다.

Denosing Algorithm

위 정의에 따라 X 는 i.i.d sample of P_X 이다,

Structure on the sample-based graph

- Sample X로 부터 Diffusion process를 정의한다.
- 여기에서 diffusion process의 Generator, 즉, the Graph Laplacian을 유도한다.
- Graph vertices 를 X_i로 정의하고, k-nn distance $\{h(X_i)\}_{i=1}^n$ 일때, **the weight of the k-NN graph** 는 다음과 같이 정의한다.

if $\|X_i - X_j\| < \max\{h(X_i), h(X_j)\}$

$$w(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{(\max\{h(X_i), h(X_j)\})^2}\right), \tag{4}$$

Otherwise $w(X_i, X_j) = 0$ 또한 graph에 Loop가 없어도 0이다.

- **The degree function** d

$$d(X_i) = \sum_{j=1}^n w(X_i, X_j) \tag{5}$$

-**The inner production (or Riemmanian metric)** between two hilbert space $\mathcal{H}_V, \mathcal{H}_E$ (V denotes vertices, E denotes Edges)

$$\langle f, g \rangle_{\mathcal{H}_V} = \sum_{i=1}^n f(X_i) g(X_i) d(X_i), \quad \langle \phi, \psi \rangle_{\mathcal{H}_E} = \sum_{i,j=1}^n w(X_i, X_j) \phi(X_i, X_j) \psi(X_i, X_j) \tag{6}$$

- **The discerete differential**

$$\nabla : \mathcal{H}_V \to \mathcal{H}_E, \quad (\nabla f)(X_i, X_j) = f(X_i) - f(X_j) \tag{7}$$

- **the Graph Laplacian**

$$\Delta : \mathcal{H}_V \to \mathcal{H}_V, \Delta = \nabla * \nabla, \quad (\Delta f)(X_i) = f(X_i) - \frac{1}{d(X_i)} \sum_{j=1} w(X_i, X_j) f(X_j) \tag{8}$$

- Defining the matrix D with the degree function on the **diagonal** the graph Laplacian in matrix form (see [2])

$$D = I - D^{-1} W \tag{9}$$

위 방정식은 Graph Laplacian 정의를 Matrix 형태로 바꾸었을 뿐이다. 즉,

$$\Delta f = Df = (I - D^{-1}W)f = f - D^{-1}Wf \tag{10}$$

The denoising algorithm

Graph Laplacian이 Graph 상에서의 Diffusion Process의 Generator 이므로다음과 같이, 그래프상에서의 미분 방정식을 정의한다.

$$\partial_t X = -\gamma \Delta X \tag{11}$$

where $\gamma > 0$ is the diffusion constant

- By the Implicit Euler-scheme, the above equation is

$$X(t+1) - X(t) = -\delta t \gamma \Delta X(t+1) \tag{2}$$

- The solution of the implicit Euler scheme for one time step can be computed as :

$$X(t+1) = (1 + \delta t \Delta)^{-1} X(t) \tag{12}$$

- proof

$$\begin{aligned} X(t+1) + \delta t \gamma \Delta X(t+1) &= X(t) \\ (1 + \delta t \gamma \Delta) X(t+1) &= X(t) \\ X(t+1) &= (1 + \delta t \gamma \Delta)^{-1} X(t) \end{aligned}$$

- **Manifold Denoising Algorithm**
  - Choose $\delta t, k$
  - **while** Stopping Criterion not satisied **do**
    - Compute the k-NN distances $h(X_i),\ i = 1, \cdots, n$,
    - Compute the weights $w(X_i, X_j)$ of the graph with $w(X_i, X_j) = 0$

$$w(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{(\max\{h(X_i), h(X_j)\})^2}\right), \quad \text{if } \|X_i - X_j\| < \max\{h(X_i), h(X_j)\}, \tag{13}$$

    - Compute the graph Laplacian $\Delta,\ \Delta = 1 - D^{-1}W$,
    - Solve $X(t+1) - X(t) = -\delta t \gamma \Delta X(t+1) \Rightarrow X(t+1) = (1 + \delta t \Delta)^{-1} X(t)$.
  - **end while**

Diffusion and Tikonov regularization

위에서 나온 The solution of the implicit Euler scheme 은 다음의 Regularization problem on the graph의 solution과 동등하다.

$$\underset{Z^\alpha \in \mathcal{H}_V}{\arg\min} S(Z^\alpha) := \underset{Z^\alpha \in \mathcal{H}_V}{\arg\min} \sum_{\alpha=1}^{d} \|Z^\alpha - X^\alpha(t)\|^2_{\mathcal{H}_V} + \delta t \sum_{\alpha=1}^{d} \|\nabla Z^\alpha\|^2_{\mathcal{H}_V} \tag{14}$$

where Z^α denotes the α-component of the vector $Z \in \mathbb{R}^d$, and $\|\nabla Z^\alpha\|^2_{\mathcal{H}_V} = \langle Z^\alpha, \Delta Z^\alpha \rangle$.

위 방정식을 미분하면

$$\frac{\partial S(Z^\alpha)}{\partial Z^\alpha} = 2(Z^\alpha - X^\alpha(t)) + 2\delta t \Delta Z^\alpha = 0, \quad \alpha = 1, \cdots, d. \tag{15}$$

따라서

$$Z = (1 + \delta t \Delta)^{-1} X_t \tag{16}$$

그러므로 Diffusion Process의 매 스텝은 Regression Problem으로 보이게 되고 새로운 Step Z는 $X(t)$에 대한 Filtering된 결과임.

K-Nearest neighbor graph versus h-neighborhood graph

- K-NN을 사용하는 이유는 K-NN을 통해 얻어진 Graph가 h-Neighborhood 방법으로 얻어진 Graph 보다 좋은 3가지 특징이 있기 때문이다.
    - Graph가 더 좋은 Connectivity를 가지고 있다.
        - Data의 Density 차이로 인해 h-Neighborhood 에서 끊어지거나 가깝게 붙을 수 있는 graph가 k-NN 그래프에서는 더 좋으 ㄴ결과를 보여준다.
        - 매우 큰 Noise 환경에서 강인하다.
        - 변수 k의 조절에 따라 Weight matrix W와 Laplacian Δ의 sparsity를 조절하기 쉽다.

Stopping Criterion

- Diffusion이 너무 오랫동안 이루어져 데이터가 Disconnected 되거나 하나의 클러스터로 몰리는 경우
- 또 하나의 경우는 Intrinsic Dimension에 대한 정보를 선험적으로 가지고 있어 sample의 Estimated dimension 이 Intrinsic Dimension과 같을 때.

알고리즘을 Stop 시킨다.

Large sample limit and theoritical analysis

다음에 나오는 Theorem은 다음 논문 [2] 에서 유도된 결론이다.

즉, **Graph Laplacian**에 대한 내용이다.

Theorem 1

Let $\{X_i\}_{i=1}^n$ be an i.i.d. samples of a probablity measure P_M on a m-dimensional compact submanifold M of $\mathbb{R}^d$ has a density $p_M \in C^3(M)$. Let $f \in C^3(M)$ and $x \in M \setminus \partial M$, then if $g \rightarrow 0$ and $nh^{m+2} / \log n \rightarrow \infty$,

$$\lim_{h \rightarrow 0} \frac{1}{h^2}(\Delta f)(x) \sim -(\Delta_M f)(x) - \frac{2}{p} \langle \nabla f, \nabla p \rangle_{T_x M} \tag{17}$$

where Δ_M is the **Laplace-Beltrami** operator of M and $\sim$ means upto a constant which depends on the kernel function $k(\|x-y\|)$ used to define the weights $W(x,y) = k(\|x-y\|)$ of the graph.

Note.

위 방정식 자체는 논문 [2]에서 Definition으로 주어졌다 (pp.1334). 그러나. 실제 논문은 A. Grigoryan [5] 의 논문을 찾아서 유도해야 한다. [2]에서 일부 유도 방법을 해설하였으나, 실제로는 [5]에서 제대로 유도를 해야한다.

The noise-free case

Noise free case에서의 경우 식 (2)를 통해 다음과 같이 간단히 유도한다.

$$\frac{i(t+1) - i(t)}{\delta t} = -\frac{1}{\delta t} \Delta i = -\frac{h^2}{\delta t} \frac{1}{h^2} \Delta i \tag{18}$$

이때, **Diffusion Constant** $D = \frac{h^2}{\delta t}$ 으로 놓고 $h \rightarrow 0$ and $\delta t \rightarrow 0$ 에서 유한한 값을 가진다고 하면 **Theorem 1** 에서 다음의 Differential Equation을 유도할 수 있다.

$$\lim_{h\to 0} -D\frac{1}{h^2}\Delta i = D[\Delta_M i + \frac{2}{p}\langle \nabla p, \nabla i\rangle_{T_xM}] = \lim_{\delta t\to 0}\frac{i(t+1)-i(t)}{\delta t} = \partial_t i \tag{19}$$

이때, k-NN graph에서 neighborhood size h 는 local density의 함수로서 Diffusion constant D가 local density $p(x)$의 함수가 되도록 한다. $D = D(p(x))$

Lemma 2

Let $i : M \to \mathbb{R}$ be a regular, smooth embedding of an m-dimensional manifold M, then $\Delta_M i = mH$ where H is the **mean curvature of M**

이것은 매우 중요하고 독특한 Lemma 이다.

함수 i가 M위에서 regular smooth embedding 이면 이것의 **Laplace-Beltrami Operation의 결과는 mean Curvature가 된다**는 것. 그러므로 Noise free 인 경우의 Diffusion방정식

$$\partial_t i = D[\Delta_M i + \frac{2}{p}\langle \nabla p, \nabla i\rangle_{T_xM}] \tag{4}$$

은 $\Delta_M i = mH$ 에서 다음과 같이 변한다.

$$\partial_t i = D[mH + \frac{2}{p}\langle \nabla p, \nabla i\rangle_{T_xM}] \tag{5}$$

이는 M위에 Probability measurer가 Constant 라면 다음과 동등하다.

$$\partial_t i = mH \tag{20}$$

다시말해, mean Curvature flow는 Manifold에서 Denosing Effect와 동등하다. 하지만, Lemma의 결과식 (4), (5)에서 보듯이 M위에 Non-uniform Probability Measure가 존재하면 $\nabla p(x) \neq 0$ 이므로 Additional term이 존재하게 된다. 이것이 Noise Case 해석과 연관된다.

The noisy case

The Large sample, 즉, $n \to \infty$의 경우 그 때의 graph Laplacian Δ at a sample point X_i 는 다음과 같이 주어진다.

$$\Delta X_i = X_i - \frac{\int_{\mathbb{R}^d} k_h(\|X_i - y\|) y p_X(y) dy}{\int_{\mathbb{R}^d} k_h(\|X_i - y\|) p_X(y) dy} \tag{6}$$

where $k_h(\|x-y\|)$ is the weight function used in the construction of the graph.

본 논문의 경우

$$k_h(\|x-y\|) = e^{-\frac{\|x-y\|^2}{2h^2}} 1_{\|x-y\|\leq h} \tag{21}$$

그런데 이것과 비슷한 kernel은 SOFM에서도 사용하고 있다.

그리고 다음의 3가지를 가정한다.

1. The noise level σ is small compared to the neighborhood size h
2. The curvature of M is small compared to h
3. The density p_M varies slowly along M

이 경우 $-\Delta X_i$는 X_i에서 p_X의 Gradient Direction이 된다.

이 효과를 Noise free 경우에서의 mean curvature part와 분리하여 생각해 본다면, 먼저, 방정식 (6)에서의 p_X가 Gaussian이라는 가정하에 살펴보면

$$\begin{aligned}\int_{\mathbb{R}^d} k_h(\|X-y\|) y p_X(y) dy &= \int_M \frac{1}{(2\pi\sigma^2)^{d/2}} \int_{\mathbb{R}^d} k_h(\|X-y\|) y e^{-\frac{\|y-i(\theta)\|^2}{2\sigma^2}} p(\theta) dy dV(\theta) \\ &= \int_M K_h(\|X-i(\theta)\|) i(\theta) p(\theta) dV(\theta) + O(\sigma^2)\end{aligned}$$

이것을 사용하여 X 에 가장 가까운 submanifold M 상의 한 점에 가장 가까운 점을

$$X : i(\theta_{\min}) = \arg \min_{i(\theta) \in M} \|X - i(\theta)\| \tag{22}$$

로 정의하면, Curvature 조건에서 diffusion step $-\Delta X$ 는 다음과 같이 정의된다.

$$-\Delta X \approx i(\theta_{\min}) - X - \left(i(\theta_{\min}) - \frac{\int_M K_h(\|X-i(\theta)\|) i(\theta) p(\theta) dV(\theta)}{\int_M K_h(\|X-i(\theta)\|) p(\theta) dV(\theta)} \right) \tag{23}$$

여기서 우측항의 괄호 안의 부분은 $-\Delta M i(\theta_{\min}) - \frac{2}{p}\langle \nabla p, \nabla i \rangle = -mH - \frac{2}{p}\langle \nabla p, \nabla i \rangle$의 approximation 이다. 반면, 그 앞의 부분은 Noise Reduction 부분이다.

Experiments

Appendix

Introduction

surface parameter $x(q)$ 를 $x \in \mathbb{R}^n$, $q \in \mathbb{R}^d$ for $n > d$ 라 정의하고 다음과 같은 metric tensor를 정의하자.

$$g_{ij} = \frac{\partial x}{\partial q_i} \cdot \frac{\partial x}{\partial q_j} \tag{24}$$

$G = [g_{ij}]$ 는 symmetruc matrix 로서 curve의 길이, surface내 patch의 면적등을 계산하기 위한 모든 Information을 가지게 된다.

The first fundamental form

이때 **The first fundamental form**을 다음과 같이 정의하자.

$$\mathcal{F}^{(1)}(dq, dq) \doteq dq^T G(q) dq \tag{25}$$

- 이를 이렇게 생각해보자. $t : \mathbb{R} \rightarrow q(t) \in \mathbb{R}^d \rightarrow x(q) \in \mathbb{R}^n$ 인 상태에서

$$\frac{dx}{dt} = \left[\frac{\partial x}{\partial q}\right] \frac{\partial q}{\partial t} \in \mathbb{R}^n \tag{26}$$

이어야 한다. 그렇다면 $\frac{\partial q}{\partial t} \in \mathbb{R}^d$ 일때, $\left[\frac{\partial x}{\partial q}\right] \in \mathbb{R}^{n \times d}$ 이다.

따라서, $\left[\frac{\partial x}{\partial q}\right]$는 일반적인 Matrix 이므로 다루기가 쉽지 않다. 여기에 Symmetry 성질을 부여한 $d \times d$ matrix를 생각하면, 다음과 같이 정의할 수 있다.

$$G(q) = \left[\frac{\partial x}{\partial q}\right]^T \left[\frac{\partial x}{\partial q}\right] \in \mathbb{R}^{d \times d} \tag{27}$$

즉, Jacobian $J(q) = \left[\frac{\partial x}{\partial q}\right]$ 에 대하여 $G(q) = J^T J$ 이다.

한편 $q = q(s)$ 라고 할 때 parameterize surface의 Coordinate Change는 Chain-rule에 의해 다음과 같다.

$$dq = J(s)ds, \quad \text{where } J(s) = \left[\frac{\partial q}{\partial s_i}, \frac{\partial q}{\partial s_j}\right] \tag{28}$$

좌표게 변환에 대하여 Fundamental form은 Invariance 이므로 $\mathcal{F}_q^{(1)} = \mathcal{F}_s^{(1)}$ i.e.

$$ds^T G_s(s) ds = ds^T J^T(s) G_q(q(s)) J(s) ds \tag{29}$$

다시말해, The metric Tensor Transform under coordinate change as

$$G_s(s) = J^T(s) G_q(q(s)) J(s) \tag{30}$$

그러므로 만일 $G(s)$를 알 수 있다면 Differential form의 주요 정보를 알 수 있다. 예를 들어, $\tilde{x}(t) = x(q(t))\text{for } t \in [t_1, t_2]$ 로 정의되는 Curve의 Arc Length를 구하는 경우

$$L(t_1, t_2) = \int_{t_1}^{t_2} \left(\frac{d\tilde{x}}{dt} \cdot \frac{d\tilde{x}}{dt}\right)^{\frac{1}{2}} dt = \int_{t_1}^{t_2} \left(\left[\frac{dq}{dt}\right]^T G(q) \frac{dq}{dt}\right)^{\frac{1}{2}} dt \tag{31}$$

Notice

$$\left(\frac{d\tilde{x}}{dt}\right)^T \left(\frac{d\tilde{x}}{dt}\right) = \left(\left[\frac{\partial x}{\partial q}\right] \frac{\partial q}{\partial t}\right)^T \left(\left[\frac{\partial x}{\partial q}\right] \frac{\partial q}{\partial t}\right) = \left(\frac{\partial q}{\partial t}\right)^T \left[\frac{\partial x}{\partial q}\right]^T \left[\frac{\partial x}{\partial q}\right] \left(\frac{\partial q}{\partial t}\right) = \left(\frac{\partial q}{\partial t}\right)^T G(q) \left(\frac{\partial q}{\partial t}\right) \tag{32}$$

또한 element of surface area는 다음과 같다.

$$dS = |G(q)|^{\frac{1}{2}} dq_1 \wedge \cdots \wedge dq_n \tag{33}$$

where $|G(q)|^{\frac{1}{2}} \doteq \sqrt{\det G(q)}$

$G = [g_{ij}]$ 로 정의되었는데, Metric Tensor의 Inverse를 다음과 같이 표시한다.

$$G^{-1} = [G^{ij}] \tag{34}$$

Gradient and Divergence : For example,

Gradient vector field of a real-valued function on the surface cna be defined as

$$(\nabla f)_i \doteq \sum_j g^{ij} \frac{\partial f}{\partial q_j} \tag{35}$$

- 보다 정확히 표시하면 다음의 의미이다.

$$\nabla_x f = \sum_i \frac{\partial f}{\partial x_i} e_i^x = \sum_i \left(\sum_j g^{ij} \frac{\partial f}{\partial q_j}\right) e_i^x \tag{36}$$

  - 즉, e^q Frame에서 만들어진 Gradient의 계수는 Metric Tensor의 Inverse g^{ij}를 통해 e^x Frame의 Gradient의 계수로 보내는 것이다.
  - 이를 통해 서로 다른 Frame 혹은 Tangent Space에서의 Gradient 를 모든 정보가 없어도 Frame간의 Metric Tensor만 있다면 구할 수 있다.

proof $f(x(q))$ 에 대하여 생각하면 간단하다. 먼저 x에 대한 Orthogonal Frame을 $e_k^x \triangleq \frac{\partial}{\partial x_k}$ 라 정의하면 q에 대한 Frame vector는

$$e_i^q \triangleq \frac{\partial}{\partial q_i} = \frac{\partial}{\partial x_k} \frac{\partial x_k}{\partial q_i} = e_k^x \frac{\partial x_k}{\partial q_i} \tag{37}$$

이때 $e_k^x \in \mathbb{R}$ 이고 $|\frac{\partial x_k}{\partial q_i}| = 1$ 이다. 만일 $e_k^x \in \mathbb{R}^n$ 이면 $|e_i^q| = 1$ 이므로

$$e_i^q \triangleq \alpha_i \frac{\partial}{\partial q_i} = \frac{1}{|J_i(q)|} \sum_k \frac{\partial}{\partial x_k} \frac{\partial x_k}{\partial q_i} = \frac{1}{|G_i(q)|^{\frac{1}{2}}} \sum_k e_k^x \frac{\partial x_k}{\partial q_i} \tag{38}$$

위 식에서 $J_i(q)$는 Jacobian의 i 번째 Column Vector를 가리키며 $J_i(q)$는 크기이다. 마찬가지로 $G_i(q)$ 가 결정된다.

Gradient는 각 Frame에 대하여 다음과 같다.

$$\nabla_x f(x) = \sum_k \frac{\partial f}{\partial x_k} e_k^x, \;\; \nabla_q f(x(q)) = \sum_i \frac{\partial f(x(q))}{\partial q_i} e_i^q \tag{39}$$

논의를 간편하게 하기 위해 $|G_i(q)| = 1$인 경우만 생각하자.

위 Gradient 식은 좌항의 e_k^x Frame상의 Gradient를 e_i^q Frame에 대하여 표현하는 것이므로

$$\begin{aligned}
\nabla_q f(x(q)) &= \sum_i \frac{\partial f}{\partial q_i} e_i^q \\
&= \sum_i \sum_k \frac{\partial f}{\partial x_k} \frac{\partial x_k}{\partial q_i} e_i^q \\
&= \sum_k \sum_{i,j} \frac{\partial f}{\partial x_k} \frac{1}{|G_i(q)|^{\frac{1}{2}}} \frac{\partial x_k}{\partial q_i} \cdot \frac{\partial x_k}{\partial q_j} e_k^x = \sum_k \sum_{i,j} \frac{\partial x_k}{\partial q_i} \frac{\partial x_k}{\partial q_j} \frac{\partial f}{\partial x_k} e_k^x \\
&= \sum_k \sum_{ij} g_{ij} \frac{\partial f}{\partial x_k} e_k^x = G(q) \nabla f(x)
\end{aligned}$$

따라서 g_{ij} 의 Inverse의 정의에 의해 ($\because G(q)^{-1} \triangleq | G(q) |^{-\frac{1}{2}} \; [\tilde{g}^{ij}] = [g^{ij}]$)

$$(\nabla_x f)_i = \frac{\partial f}{\partial x_i} = \sum_j g^{ij} \frac{\partial f}{\partial q_j} \tag{40}$$

G가 $d \times d$ matrix이므로 단순 Matrix 표현식으로는 $n \times 1$ 벡터인 $\nabla_x f$ 를 표현할 수 없으며, 내부 Component의 관계를 통해 얻어 낼 수 밖에 없다. **Q.E.D**

Example : Gradient on Sphere Coordinate

The spherical coordinate 는 다음과 같다.

$$\begin{aligned}
x &= r \sin\theta \cos\phi \\
y &= r \sin\theta \sin\phi \\
z &= r \cos\theta
\end{aligned}$$

그리고 $e_k^x \in \{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\}$ 라 하고 $e_k^\phi \in \{\alpha_1 \frac{\partial}{\partial r}, \alpha_2 \frac{\partial}{\partial \phi}, \alpha_3 \frac{\partial}{\partial \theta}\}$ 이라 하자,

let J be the Jacobi matrix

$$J = \begin{pmatrix} \sin\theta\cos\phi & -r\sin\phi\sin\theta & r\cos\phi\cos\theta \\ \sin\theta\sin\phi & r\cos\phi\sin\theta & r\sin\phi\cos\theta \\ \cos\theta & 0 & -r\sin\theta \end{pmatrix}, \tag{41}$$

the metric tensor can be obtained as

$$[g_{ij}] = J^T J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2\sin^2\theta & 0 \\ 0 & 0 & r^2 \end{pmatrix} \tag{42}$$

좌표 변환을 위해 다음을 각각 살펴보면

$$
\begin{aligned}
\frac{\partial}{\partial r} &= \frac{\partial}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial}{\partial z}\frac{\partial z}{\partial r} \\
\frac{\partial}{\partial r} &= \frac{\partial}{\partial x}\cos\phi\sin\theta + \frac{\partial}{\partial y}\sin\phi\sin\theta + \frac{\partial}{\partial z}\cos\theta \\
\left|\frac{\partial}{\partial r}\right| &= \sqrt{\cos^2\phi\sin^2\theta + \sin^2\phi\sin^2\theta + \cos^2\theta} = 1
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \phi} &= \frac{\partial}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial}{\partial y}\frac{\partial y}{\partial \phi} + \frac{\partial}{\partial z}\frac{\partial z}{\partial \phi} \\
\frac{\partial}{\partial \phi} &= \frac{\partial}{\partial x}(-r\sin\phi\sin\theta) + \frac{\partial}{\partial y}r\cos\phi\sin\theta + \frac{\partial}{\partial z}0 \\
\left|\frac{\partial}{\partial \phi}\right| &= \sqrt{r^2\cos^2\phi\sin^2\theta + \sin^2\phi\sin^2\theta} = r\sin\theta
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \theta} &= \frac{\partial}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial}{\partial z}\frac{\partial z}{\partial \theta} \\
\frac{\partial}{\partial \theta} &= \frac{\partial}{\partial x}r\cos\phi\cos\theta + \frac{\partial}{\partial y}r\sin\phi\cos\theta + \frac{\partial}{\partial z}(-r\sin\theta) \\
\left|\frac{\partial}{\partial r}\right| &= \sqrt{r^2\cos^2\phi\sin^2\theta + r^2\sin^2\phi\sin^2\theta + r^2\cos^2\theta} = r
\end{aligned}
$$

그러므로 Spherical Coordinate에서의 Element Vector e_k^ϕ는 다음과 같다.

$$
\begin{aligned}
1 = |e_1^\phi| = |\alpha_1|\left|\frac{\partial}{\partial r}\right| &\Rightarrow |\alpha_1| = 1 && \Rightarrow e_1^\phi = \frac{\partial}{\partial r} \\
|e_2^\phi| = |\alpha_2|\left|\frac{\partial}{\partial \phi}\right| &\Rightarrow |\alpha_2| = \frac{1}{r\sin\theta} && \Rightarrow e_2^\phi = \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi} \\
|e_3^\phi| = |\alpha_3|\left|\frac{\partial}{\partial \theta}\right| &\Rightarrow |\alpha_3| = \frac{1}{r} && \Rightarrow e_3^\phi = \frac{1}{r}\frac{\partial}{\partial \theta}
\end{aligned}
$$

따라서

$$
\begin{aligned}
\nabla f(x) &= \frac{\partial f}{\partial r}\frac{\partial}{\partial r} + \frac{1}{r^2\sin^2\theta}\frac{\partial f}{\partial \phi}\frac{\partial}{\partial \phi} + \frac{1}{r^2}\frac{\partial f}{\partial \theta}\frac{\partial}{\partial \theta} \\
&= \frac{\partial f}{\partial r}e_1^\phi + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}e_2^\phi + \frac{1}{r}\frac{\partial f}{\partial \theta}e_3^\phi \\
&= (\frac{\partial f}{\partial r},\ \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi},\ \frac{1}{r}\frac{\partial f}{\partial \theta})^T
\end{aligned}
$$

Divergence of a vector field on the surface

- Definition of Divergence $\nabla \cdot f : M \to \mathbb{R}$ for $f : \mathbb{R}^n \to \mathbb{R}^n$
  - $\nabla \cdot f = \mathrm{trace}(df)$
- Definiton of Divergence for Frame $\{e_k^x\}$ and a vector field $f(x) \in \mathbb{R}^n$

$$
\nabla \cdot f(x) = \sum_k \frac{\partial f_k}{\partial x_k} \tag{43}
$$

그러므로 Divergence는 다음과 같이 정의될 수 있다.

For $e_k^x \triangleq \frac{\partial}{\partial x_k}$, and $f(x) = \sum_i f_i e_i^x$

$$
\begin{aligned}
\nabla \cdot f(x) &= \sum_k \frac{\partial}{\partial x_k} e_k^x \cdot \sum_i f_i e_i^x \\
&= \sum_k \sum_i \frac{\partial f_i}{\partial x_k} e_k^x \cdot e_i^x = \sum_k \frac{\partial f_k}{\partial x_k} \quad \because e_k^x \cdot e_i^x = 1 \text{ if } k = i, \text{ else } 0
\end{aligned}
$$

만일 위에서 $e_k^x \cdot e_i^x$ Euclidean Space가 아니라면 Metric Tensor에 의한 해석이 필요하다. 이때 Divergence는 다음과 같이 정의된다.

$$
\nabla \cdot f = tr(Y \mapsto \nabla_Y f) \tag{44}
$$

이때 Vector field f 는

$$
f = f^i \frac{\partial}{\partial q_i} e_i^q \triangleq f^i \partial_i \tag{45}
$$

Divergence의 정의를 생각해보면 Trace이므로 Scalar 값이다. 이를 생각해보면 특정 Component의 값으로 만들어 줄 수 있다는 것이 된다. 일반적으로 Vector field $V = \sum_j v^j X_j$ 를 정의하고 이것의 Covariant Differential을 생각해보면

$$
\frac{DV}{dt} = \sum_k \left[\frac{dv^k}{dt} + \sum_{ij} \Gamma_{ij}^k \frac{dx^i}{dt} v_j \right] X_k \tag{46}
$$

그런데 여기서 **출력을 k가 아닌 i 성분으로 보내면 이는 Trace 중의 한 성분과 같은 의미**가 된다.

이때, Divergence는 다음과 같이 **Normal ([partial)성분의 미분** (그래서 **Tangent Space위로 Projection**)과 **Tangent 성분의 미분** (Christoffel Symbolic part)으로 표현된다.

$$
\nabla \cdot f = \frac{\partial f^i}{\partial q_i} + \Gamma_{ij}^i f^j \tag{47}
$$

그러면 k 대신 하나의 성분 i로만 보내는 것이므로 Christoffel 기호는

$$
\Gamma_{ij}^i = \frac{1}{2} \left(\frac{\partial}{\partial q_i} g_{jk} + \frac{\partial}{\partial q_j} g_{ki} - \frac{\partial}{\partial q_k} g_{ij} \right) g^{ki} \tag{48}
$$

이 경우

$$
g^{ki} \frac{\partial}{\partial q_i} g_{jk} = g^{ki} \frac{\partial}{\partial q_k} g_{ij} \tag{49}
$$

즉, 아래 위 첨자를 지워보면 g_j 만 양변에 똑같이 남게되어 같다.

이를 이용하여 다음을 증명한다.

Divergence of a vector field

Divergence of a vector field on the surface 는 따라서 다음과 같이 정의된다.

$$
\nabla \cdot f \doteq |G|^{-\frac{1}{2}} \sum_k \frac{\partial}{\partial q_k} (|G|^{\frac{1}{2}} f_k) \tag{50}
$$

proof

위에서 Trace의 경우 Christoffel 기호는

$$
\Gamma_{ij}^i = \frac{1}{2} g^{ki} \frac{\partial}{\partial q_j} g_{ki} \tag{51}
$$

Jacobi Formula의 Corollary에서

$$\frac{1}{2} g^{ki} \frac{\partial}{\partial q_j} g_{ki} = \frac{\partial}{\partial q_j} \log \sqrt{|G|} \tag{52}$$

Label을 수정하고 Divergence 정의에서

$$\begin{aligned}(\nabla \cdot f)^i &= \frac{\partial f^i}{\partial q_i} + \Gamma^i_{ij} f^j \\ &= \frac{\partial f^i}{\partial q_i} + \frac{\partial}{\partial q_i} \log \sqrt{|G|} f^i \\ &= \frac{\partial f^i}{\partial q_i} + \frac{1}{\sqrt{|G|}} \frac{\partial \sqrt{|G|}}{\partial q_i} f^i \\ &= \frac{1}{\sqrt{|G|}} \frac{\partial \sqrt{|G|} f^i}{\partial q_i}\end{aligned}$$

그러므로

$$\nabla \cdot f = \frac{1}{\sqrt{|G|}} \sum_i \frac{\partial \sqrt{|G|} f^i}{\partial q_i} \tag{53}$$

Laplace Beltrami Operator

- **The Laplace (or Laplace Beltrami Operator)** of the smooth real-valued function os defined as **the divergence of the gradient**

$$\nabla \cdot \nabla f \doteq |G|^{\frac{1}{2}} \sum_i \frac{\partial}{\partial q_i} \left(|G|^{\frac{1}{2}} \sum_j g^{ij} \frac{\partial f}{\partial q_j} \right) \tag{54}$$

위에서의 Gradient 정의와 Divergence 정의를 가져와서 Laplace -Beltrami 정의에 대입하면 증명 끝.

Jacobi Formula : Matrix Defferentiation for $\det A$

$$\frac{d}{dt} \det A(t) = \operatorname{tr} \left(\operatorname{adj} A \cdot \frac{dA(t)}{dt} \right) \tag{55}$$

Lemma 1

$$\frac{d \det I(s)}{ds} = \text{trace} \tag{56}$$

그러므로 differential $\det'(I)$ 는 Linear operator로서 $\det'(I)$: $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}$

proof

$$\frac{d}{dT} \det(I)(T) = \lim_{\varepsilon \rightarrow 0} \frac{\det(I + \varepsilon T) - \det I}{\varepsilon} \tag{57}$$

논의를 간단하게 하기 위해여 T도 Diagonal $n \times n$ matrix라고 하면 (그냥 해도 결과는 같다.)

$$\det(I + \varepsilon T) - \det I = 1 + \binom{n}{1} \varepsilon T + \cdots + \binom{n}{n} \varepsilon^n T^n - 1 \tag{58}$$

에서 ε이 1차인 항은 $n \varepsilon T$ 뿐이다. 따라서 Trace.

Note

그러므로

$$\frac{dI(T)}{dT} = \text{trace} \Rightarrow dI(T) = \text{trace T} \tag{59}$$

Lemma 2

For an invertible matrix A, we have :

$$\frac{d \det A(T)}{dT} = \det A \cdot \text{tr}(A^{-1}T) \tag{60}$$

proof

임의의 정방형 matrix X에 대하여

$$\det X = \det(AA^{-1}X) = \det(A)\det(A^{-1}X) \tag{61}$$

이므로

에 대하여 미분하고 이 결과를 $X = A$로 대입하면

$$\begin{aligned}
\left.\frac{d \det X(T)}{dT}\right|_{X=A} &= \left.\frac{d}{dT}\det\left(AA^{-1}X\right)(T)\right|_{X=A} \\
&= \det(A)\frac{d}{dT}\det(A^{-1}A)(T) \\
&= \det(A)\frac{d \det I}{dA^{-1}T}\frac{dA^{-1}T}{dT}T \\
&= \det(A)\frac{d \det I}{dA^{-1}T}\left(A^{-1}T\right) \\
&= \det(A)\text{tr}\left(A^{-1}T\right) \quad \because \text{by Lemma 1}
\end{aligned}$$

Theorem : Jacobi's Lemma

$$\frac{d}{dt}\det A(t) = \text{tr}\left(\text{adj}A \cdot \frac{dA(t)}{dt}\right) \tag{62}$$

where **adj** is the adjugateof A such that

$$A^{-1} = \frac{\text{adj}A}{\det A}, \quad \text{where adj}A \in \mathbb{R}^{n\times n} \tag{63}$$

proof

In Lemma 2, Let $T = \frac{dA(t)}{dt}$, then

$$\frac{d}{dt}\det A(t) = \det A \cdot \text{tr}\left(A^{-1}\frac{dA(t)}{dt}\right) = \det A \cdot \text{tr}\left(\frac{\text{adj}A}{\det A}\cdot\frac{dA}{dt}\right) = \text{tr}\left(\text{adj A}\cdot\frac{dA}{dt}\right) \tag{64}$$

Corollary

Let $U \subseteq \mathbb{R}^n$ be openm and let $g : U \rightarrow GL(n, \mathbb{C}) \subseteq M_n(\mathbb{C})$ be differentible. g의 component를 g_{ij} 로 표현하고 그 Inverse를 g^{ij} 라 할 때

$$g^{ij}\frac{\partial g_{ij}}{\partial q_k} = \frac{\partial g_{ij}g^{ij}}{\partial q_k} = \frac{1}{\det g}\frac{\partial \det g}{\partial q_k} = \frac{\partial}{\partial q_k}\log(|\det g|) \tag{65}$$

위 표현은 Manifold 의 Tangent space의 vector space를 $\frac{\partial}{\partial q_k}$로 보았을 때 이다. 이를 단순하게 X_k 로 보고 이를 생략한 형태로 다시 쓰면 다음과 같다.

$$g^{ij}\partial_k g_{ij} = \partial_k g_{ij} g^{ij} = \frac{\partial_k \det g}{\det g} = \partial_k \log(|\det g|) \tag{66}$$

proof

Fix $x \in U$ and consider the family $h(t) = g(x + tq_k)$ then,

$$\left.\frac{\partial h(t)}{\partial t}\right|_{t=0} = \frac{\partial g(x)}{\partial q_k} \tag{67}$$

and

$$\left.\frac{\partial \det h(t)}{\partial t}\right|_{t=0} = \frac{\partial(\det g(x))}{\partial q_k} \tag{68}$$

By Jacobi Formula

$$\left.\frac{\partial \det h(t)}{\partial t}\right|_{t=0} = \mathrm{tr}\left(\mathrm{Adj}(h(0))\left.\frac{\partial h(t)}{\partial t}\right|_{t=0}\right) = \det(g(x))\mathrm{tr}\left(g^{-1}(x)\frac{\partial g(x)}{\partial q_k}\right) = \det(g(x))g^{ij}\frac{\partial g_{ij}}{\partial q_k} \tag{69}$$

그러므로

$$\frac{1}{\det(g(x))}\frac{\partial(\det g(x))}{\partial q_k} = \frac{\partial}{\partial q_k}\log(|\det g(x)|) \tag{70}$$

그리고 나머지는 모두 성립한다.

Levi-Civita Connection[3]

Affine Connection ∇ 에서 다음의 2가지 조건이 만족되면 이를 Levi-Civita Connection이라 한다.

1. **it preserves the metric**, i.e., $\nabla g = 0$.
2. **it is torsion-free**, i.e., for any vector fields X and Y we have $\nabla_X Y - \nabla_Y X = [X, Y]$, where $[X, Y]$ is the Lie bracket of the vector fields X and Y.

Christoffel synbol

한편 Christoffel 기호는 다음과 같이 정의된다[3].

$$\Gamma^m_{ij} = \frac{1}{2}\sum_k \left(\frac{\partial}{\partial q_i} g_{jk} + \frac{\partial}{\partial q_j} g_{ki} - \frac{\partial}{\partial q_k} g_{ij}\right) g^{km} \tag{71}$$

- $ijk \rightarrow jki \rightarrow -kij$ **with** k and **out** m 으로 외우면 된다.

Note

1. Upper Index는 Scalar 의 index, Lower index는 vector의 index로 생각하면 된다.

Reference

[1] Matthias Hein, Markus Maier, "Manifold Denosing",

[2] Hein, Matthias, Audibert, Jean-Yves von Luxburg, Ulrike, "From Graphs to Manifolds – Weak and Strong Pointwise Consistency of Graph Laplacians"

[3] Do Carmo, "Riemannian Geometry", pp. 50-54

[4] G.S. Chrikijian, 'Stochastic Models, Information theory, and Lie Groups', Vol 1, Birkhauser, 2000

[5] A. Grigoryan, 'Heat kernels on weighted manifolds and applications', Cont. Math, 398:93-191, 2006