谱嵌入与拓扑结构

刘家锋

哈尔滨工业大学

谱嵌入与拓扑结构

谱嵌入方法

线性方法

- **子空间与线性流形**
  - 子空间：令W为线性空间V的非空子集，如果W满足
    1. $\forall \mathbf{x}, \mathbf{y} \in W$，有$\mathbf{x} + \mathbf{y} \in W$
    2. $\forall \mathbf{x} \in W$，$\alpha \in R$, 有$\alpha \mathbf{x} \in W$

    则称W为V的线性子空间
  - 线性流形：若W为V的线性子空间，$\boldsymbol{\mu} \in V$，则$W + \boldsymbol{\mu}$称为V的线性子流形
- **子空间学习**
  - 假设样本$\mathbf{x}$分布在一个嵌入在R^m中的m'维子空间
  - m'维子空间由基矢量$\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{m'})$所张成
  - 样本$\mathbf{x}$在m'维子空间上的投影：

$$\mathbf{y} = \mathbf{W}^t \mathbf{x}$$

主成分分析

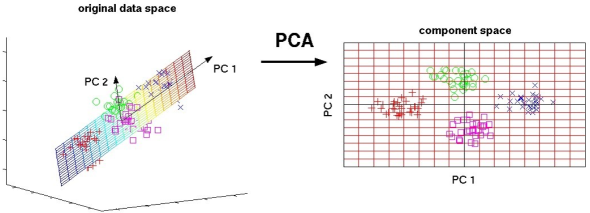
Algorithm 1 PCA: Principal Component Analysis

1: 输入样本集合D，计算均值矢量$\boldsymbol{\mu}$和协方差矩阵$\boldsymbol{\Sigma}$

2: 计算矩阵$\boldsymbol{\Sigma}$的特征值和特征矢量，特征值由大到小排序

3: 选择前m'个特征矢量作为列矢量，构成矩阵$\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{m'})$

4: m维特征矢量$\mathbf{x}$可以转换为m'维矢量：$\mathbf{y} = \mathbf{W}^t(\mathbf{x} - \boldsymbol{\mu})$



线性判别分析

Algorithm 2 LDA: Linear Discriminant Anaylsis

1: 计算类间和类内散布矩阵$\mathbf{S}_b$和$\mathbf{S}_w$

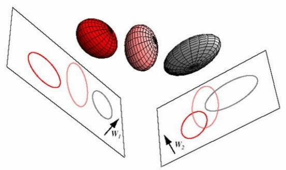
2: 计算矩阵$\mathbf{S}_w^{-1}\mathbf{S}_b$的特征值和特征矢量，特征值由大到小排序

3: 选择前m'个特征矢量作为列矢量，构成矩阵$\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{m'})$

4: m维特征矢量$\mathbf{x}$可以转换为m'维矢量：$\mathbf{y} = \mathbf{W}^t\mathbf{x}$

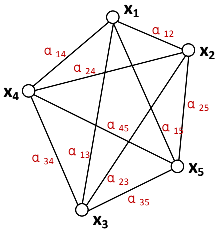
$$\mathbf{S}_w = \sum_{i}^{c} \sum_{\mathbf{x}\in D_i} (\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^t$$

$$\mathbf{S}_b = \sum_{i=1}^{c} n_i(\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^t$$



流形学习的谱方法

- **流形学习的谱嵌入框架**
  - 很多流形学习方法可以采用统一的框架来描述和求解
  - 在统一的框架下，可以设计出新的流形学习方法
- **基本思路**
  - 计算R^m空间样本$\mathbf{x}_i, \mathbf{x}_j$之间的相似度$a_{ij}$，构造相似图
  - 保持降维到$R^{m'}$维空间之后，样本$\mathbf{y}_i, \mathbf{y}_j$之间的相似度
  - 同时，满足与问题相关的某种约束条件



谱嵌入框架(Spectral Embedding)

降维之后的样本矩阵：

$$\mathbf{Y} = (\mathbf{y}_1, \cdots, \mathbf{y}_n) = \begin{pmatrix} \tilde{\mathbf{y}}_1^t \\ \vdots \\ \tilde{\mathbf{y}}_{m'}^t \end{pmatrix} = \begin{pmatrix} y_{11} & \cdots & y_{n1} \\ \vdots & \ddots & \vdots \\ y_{1m'} & \cdots & y_{nm'} \end{pmatrix}$$

其中，矢量$\tilde{\mathbf{y}}_k$由所有降维样本第k维特征构成。

相似图的相似矩阵和Laplacian矩阵：

$$\mathbf{A} = (a_{ij}), \qquad \mathbf{L} = \mathbf{D} - \mathbf{A}$$

谱嵌入优化目标

保持降维之后样本之间的相似性：

$$\min_{\mathbf{Y}} \frac{1}{2} \sum_{i \neq j} a_{ij} ||\mathbf{y}_i - \mathbf{y}_j||^2$$

目标函数转化为：

$$\begin{aligned}
\sum_{i \neq j} a_{ij} ||\mathbf{y}_i - \mathbf{y}_j||^2 &= \sum_{i \neq j} a_{ij} \sum_{k=1}^{m'} (y_{ik} - y_{jk})^2 \\
&= \sum_{k=1}^{m'} \sum_{i \neq j} a_{ij} (y_{ik} - y_{jk})^2 \\
&= \sum_{k=1}^{m'} \tilde{\mathbf{y}}_k^t \mathbf{L} \tilde{\mathbf{y}}_k
\end{aligned}$$

谱嵌入框架(Spectral Embedding)

谱嵌入的优化问题

$$\min_{\mathbf{Y}} \sum_{k=1}^{m'} \tilde{\mathbf{y}}_k^t \mathbf{L} \tilde{\mathbf{y}}_k$$

subject to

$$\tilde{\mathbf{y}}_k^t \mathbf{B} \tilde{\mathbf{y}}_k = d, \quad k = 1, \cdots, m'$$

具体的流形学习方法采用不同的相似图的Laplacian矩阵$\mathbf{L}$，不同的约束矩阵$\mathbf{B}$

子空间学习的谱嵌入

构造线性降维映射：

$$\mathbf{y}_i = \mathbf{W}^t\mathbf{x}_i, \quad i = 1, \cdots, n$$

$\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_{m'})$为变换矩阵，构造样本矩阵$\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$

$$\mathbf{Y} = (\mathbf{y}_1, \cdots, \mathbf{y}_n) = (\mathbf{W}^t\mathbf{x}_1, \cdots, \mathbf{W}^t\mathbf{x}_n) = \mathbf{W}^t\mathbf{X}$$

同时

$$\mathbf{Y} = \begin{pmatrix} \tilde{\mathbf{y}}_1^t \\ \vdots \\ \tilde{\mathbf{y}}_{m'}^t \end{pmatrix} = \begin{pmatrix} \mathbf{w}_1^t \\ \vdots \\ \mathbf{w}_{m'}^t \end{pmatrix} \mathbf{X} = \begin{pmatrix} \mathbf{w}_1^t\mathbf{X} \\ \vdots \\ \mathbf{w}_{m'}^t\mathbf{X} \end{pmatrix}$$

因此：

$$\tilde{\mathbf{y}}_k = \mathbf{X}^t\mathbf{w}_k, \quad k = 1, \cdots, m'$$

子空间学习的谱嵌入

子空间学习谱嵌入的优化问题：

$$\min_{\mathbf{W}} \quad \sum_{k=1}^{m'} \tilde{\mathbf{y}}_k^t \mathbf{L} \tilde{\mathbf{y}}_k = \sum_{k=1}^{m'} \mathbf{w}_k^t \mathbf{X} \mathbf{L} \mathbf{X}^t \mathbf{w}_k$$

subject to

$$\mathbf{w}_k^t \mathbf{X} \mathbf{B} \mathbf{X}^t \mathbf{w}_k = d, \quad k = 1, \cdots, m'$$

问题的解为广义特征值问题

$$(\mathbf{X}\mathbf{L}\mathbf{X}^t)\mathbf{w} = \lambda(\mathbf{X}\mathbf{B}\mathbf{X}^t)\mathbf{w}$$

最小m'个特征值对应的特征矢量

谱嵌入框架的意义

- **从谱嵌入的角度理解流形学习方法**
  - 在谱嵌入的框架下分析传统的流形学习方法
  - 考察传统方法对应的相似图和约束图
- **设计新的流形学习方法**
  - 根据需要，可以设计不同的相似图构造方法，得到新的流形学习方法
  - 不同相似性度量，不同的图结构对应不同的相似图

谱嵌入框架下的PCA

PCA的优化问题：将基矢量的最大值优化转换为最小值优化

$$\min_{\mathbf{W}} \quad \sum_{k=1}^{m'} \mathbf{w}_k^t \mathbf{\Sigma} \mathbf{w}_k$$

subject to

$$\mathbf{w}_k^t \mathbf{w}_k = 1, \quad k = 1, \cdots, m'$$

定义样本矩阵$\mathbf{X}$和矢量$\mathbf{e}$:

$$\mathbf{X} = (\mathbf{x}_1, \cdots, \mathbf{x}_n), \qquad \mathbf{e} = (1, \cdots, 1)^t$$

样本的均值可以表示为:

$$\boldsymbol{\mu} = \frac{1}{n}\mathbf{X}\mathbf{e} = \frac{1}{n}(\mathbf{x}_1, \cdots, \mathbf{x}_n)\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i$$

谱嵌入框架下的PCA

协方差矩阵可以写成:

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^t \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^t - \boldsymbol{\mu}\boldsymbol{\mu}^t \\
&= \frac{1}{n}(\mathbf{x}_1, \cdots, \mathbf{x}_n)\begin{pmatrix} \mathbf{x}_1^t \\ \vdots \\ \mathbf{x}_n^t \end{pmatrix} - \left(\frac{1}{n}\mathbf{X}\mathbf{e}\right)\left(\frac{1}{n}\mathbf{X}\mathbf{e}\right)^t \\
&= \frac{1}{n}\mathbf{X}\mathbf{X}^t - \frac{1}{n^2}\mathbf{X}\mathbf{e}\mathbf{e}^t\mathbf{X}^t \\
&= \frac{1}{n}\mathbf{X}(\mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}^t)\mathbf{X}^t
\end{aligned}
$$

谱嵌入框架下的PCA

如果令：

$$\mathbf{L} = \mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}^t$$

PCA的优化问题可以表示为谱嵌入框架一致的形式：

$$\min_{\mathbf{W}} \sum_{k=1}^{m'} \mathbf{w}_k^t \mathbf{X}\mathbf{L}\mathbf{X}^t \mathbf{w}_k$$

subject to

$$\mathbf{w}_k^t \mathbf{w}_k = 1, \quad k = 1, \cdots, m'$$

PCA的相似图

令：

$$\mathbf{A} = \frac{1}{n}\mathbf{e}\mathbf{e}^t = \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$$

对角矩阵：

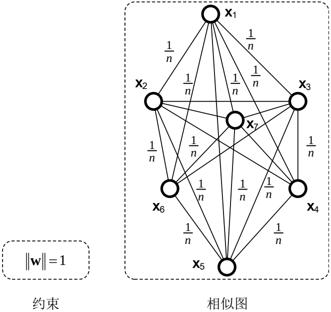
$$\mathbf{D} = \mathrm{diag}\left(\sum_{j=1}^{n} a_{ij}\right) = \mathbf{I}$$

可见矩阵**L**为Laplacian矩阵：

$$\mathbf{L} = \mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}^t = \mathbf{D} - \mathbf{A}$$

PCA的相似图

矩阵$\mathbf{A}$为PCA的相似矩阵，对应的相似图为全连接图，任意两个样本之间的相似度为$\frac{1}{n}$



谱嵌入框架下的LDA

LDA的优化目标是Rayleigh商：

$$\min_{\mathbf{w}} \frac{\mathbf{w}^t\mathbf{S}_w\mathbf{w}}{\mathbf{w}^t\mathbf{S}_b\mathbf{w}}$$

可以证明，类间散布$\mathbf{S}_b$与类内散布$\mathbf{S}_w$之间满足：

$$\mathbf{S}_b = \sum_{j=1}^{c} n_j(\boldsymbol{\mu}_j - \boldsymbol{\mu})(\boldsymbol{\mu}_j - \boldsymbol{\mu})^t = n\boldsymbol{\Sigma} - \mathbf{S}_w$$

因此：

$$\frac{\mathbf{w}^t\mathbf{S}_b\mathbf{w}}{\mathbf{w}^t\mathbf{S}_w\mathbf{w}} = \frac{n\mathbf{w}^t\boldsymbol{\Sigma}\mathbf{w} - \mathbf{w}^t\mathbf{S}_w\mathbf{w}}{\mathbf{w}^t\mathbf{S}_w\mathbf{w}} = n\frac{\mathbf{w}^t\boldsymbol{\Sigma}\mathbf{w}}{\mathbf{w}^t\mathbf{S}_w\mathbf{w}} - 1$$

Rayleigh商的优化等价于：

$$\min_{\mathbf{w}} \frac{\mathbf{w}^t\mathbf{S}_w\mathbf{w}}{\mathbf{w}^t\mathbf{S}_b\mathbf{w}} \iff \min_{\mathbf{w}} \frac{\mathbf{w}^t\mathbf{S}_w\mathbf{w}}{\mathbf{w}^t\boldsymbol{\Sigma}\mathbf{w}}$$

谱嵌入框架下的LDA

Rayleigh商优化转化为条件优化问题：

$$\min_{\mathbf{W}} \quad \sum_{k=1}^{m'} \mathbf{w}_k^t \mathbf{S}_w \mathbf{w}_k$$

subject to

$$\mathbf{w}_k^t \boldsymbol{\Sigma} \mathbf{w}_k = d, \quad k = 1, \cdots, m'$$

谱嵌入框架下的LDA

定义n维指示矢量$\mathbf{e}_j$，第i维元素指示样本$\mathbf{x}_i$是否属于类别ω_j：

$$e_{ji} = \begin{cases} 1, & \mathbf{x}_i \in \omega_j \\ 0, & \text{otherwise} \end{cases}$$

样本$\mathbf{x}_i$属于且仅属于$\omega_1, \cdots, \omega_c$中的一个类别，因此：

$$\sum_{j=1}^{c} e_{ji} = 1$$

类别ω_j中的样本数：

$$n_j = \sum_{i=1}^{n} e_{ji}$$

谱嵌入框架下的LDA

ω_j类的样本均值：

$$\boldsymbol{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n} e_{ji}\mathbf{x}_i = \frac{1}{n_j}(\mathbf{x}_1, \cdots, \mathbf{x}_n)\,\mathbf{e}_j = \frac{1}{n_j}\mathbf{X}\mathbf{e}_j$$

总的类内散布矩阵：

$$\begin{aligned}
\mathbf{S}_w &= \sum_{j=1}^{c}\sum_{i=1}^{n} e_{ji}(\mathbf{x}_i - \boldsymbol{\mu}_j)(\mathbf{x}_i - \boldsymbol{\mu}_j)^t \\
&= \sum_{j=1}^{c}\left[\sum_{i=1}^{n} e_{ji}\mathbf{x}_i\mathbf{x}_i^t - \sum_{i=1}^{n} e_{ji}\boldsymbol{\mu}_j\boldsymbol{\mu}_j^t\right] \\
&= \sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^t - \sum_{j=1}^{c} n_j\boldsymbol{\mu}_j\boldsymbol{\mu}_j^t \\
&= \mathbf{X}\mathbf{X}^t - \sum_{j=1}^{c}\frac{1}{n_j}(\mathbf{X}\mathbf{e}_j)(\mathbf{X}\mathbf{e}_j)^t = \mathbf{X}\left(\mathbf{I} - \sum_{j=1}^{c}\frac{1}{n_j}\mathbf{e}_j\mathbf{e}_j^t\right)\mathbf{X}^t
\end{aligned}$$

谱嵌入框架下的LDA

定义矩阵：

$$\mathbf{L}_{\mathsf{LDA}} = \mathbf{I} - \sum_{j=1}^{c} \frac{1}{n_j} \mathbf{e}_j \mathbf{e}_j^t, \quad \mathbf{L}_{\mathsf{PCA}} = \mathbf{I} - \frac{1}{n} \mathbf{e}\mathbf{e}^t$$

LDA的优化问题可以表示为谱嵌入框架一致的形式：

$$\min_{\mathbf{W}} \quad \sum_{k=1}^{m'} \mathbf{w}_k^t \mathbf{X} \mathbf{L}_{\mathsf{LDA}} \mathbf{X}^t \mathbf{w}_k$$

subject to

$$\mathbf{w}_k^t \mathbf{X} \mathbf{L}_{\mathsf{PCA}} \mathbf{X}^t \mathbf{w}_k = d, \quad k = 1, \cdots, m'$$

LDA的相似图

令：

$$\mathbf{A}_{\mathsf{LDA}} = \sum_{j=1}^{c} \frac{1}{n_j} \mathbf{e}_j \mathbf{e}_j^t, \quad \text{其中元素：} a_{ik} = \sum_{j=1}^{c} \frac{1}{n_j} e_{ji} e_{jk}$$

矩阵$\mathbf{A}_{\mathsf{LDA}}$的第$i$行求和：

$$d_i = \sum_{k=1}^{n} a_{ik} = \sum_{k=1}^{n} \sum_{j=1}^{c} \frac{1}{n_j} e_{ji} e_{jk} = \sum_{j=1}^{c} \frac{1}{n_j} e_{ji} \sum_{k=1}^{n} e_{jk} = \sum_{j=1}^{c} e_{ji} = 1$$

因此：

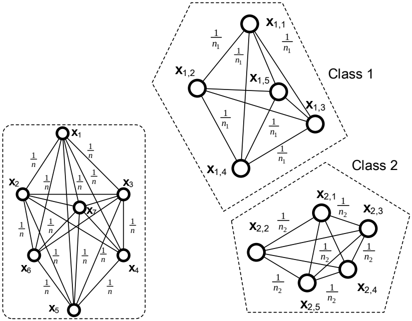
$$\mathbf{D}_{\mathsf{LDA}} = \mathrm{diag}(d_i) = \mathbf{I}$$

可见矩阵$\mathbf{L}_{\mathsf{LDA}}$为Laplacian矩阵：

$$\mathbf{L}_{\mathsf{LDA}} = \mathbf{I} - \sum_{j=1}^{c} \frac{1}{n_j} \mathbf{e}_j \mathbf{e}_j^t = \mathbf{D}_{\mathsf{LDA}} - \mathbf{A}_{\mathsf{LDA}}$$

LDA的相似图

$\mathbf{A}_{\mathrm{LDA}}$为LDA的相似矩阵，对应的相似图为分块全连接图，$\mathbf{A}_{\mathrm{PCA}}$为约束图的相似矩阵



约束 相似图

拓扑结构与近邻图

拓扑结构和相似图

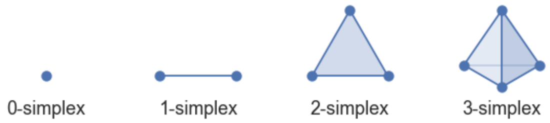
- **相似图与流形的拓扑结构**
  - 在谱嵌入框架下，数据集的相似图及其Laplacian矩阵是学习流形拓扑结构的重要手段
  - 相似图是否能够描述流形的拓扑结构？
  - 借助相似图的流形学习方法，拓扑学的理论依据是什么？

单纯形与单纯复形

- **单纯形(Simplex)**

  单纯形是对三角形、四面体的泛化，也称为超四面体；一个k维单纯形(k-simplex)由$k+1$个顶点(0-face)，$\frac{k(k+1)}{2}$个边(1-face)，以及一系列$\binom{k+1}{i+1}$个i-face 构成

  给定空间中的$k+1$个点$[x_0, \cdots, x_k]$，包含这些点的最小凸集即为一个k维单纯形

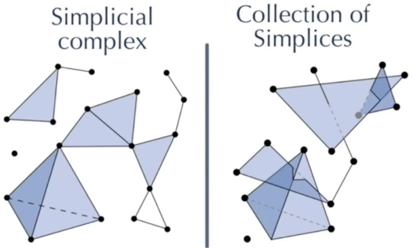


单纯形与单纯复形

- **单纯复形(Simplicial Complex)**

  单纯复形是若干个单纯形构成的集合，需要满足：

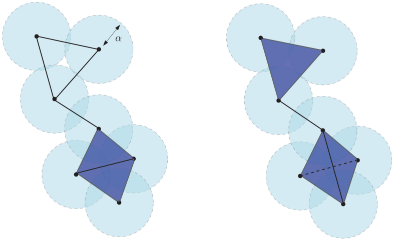
  1. 集合中每个单纯形的face都是集合的元素
  2. 任意两个单纯形的交，都是两个单纯形的face



度量空间中的单纯复形

- **Cech complex**

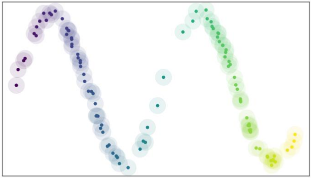
  令(X,d)是度量空间，$B(x,\alpha)$是以$x \in X$为中心，$\alpha \geq 0$为半径的球形邻域；点集$D \subset X$上的Cech复形$\mathrm{Cech}_\alpha(D)$定义为单纯形$[x_0,\cdots,x_k]$的集合，单纯形所有顶点上的$k+1$个$B(x_i,\alpha)$的交集非空

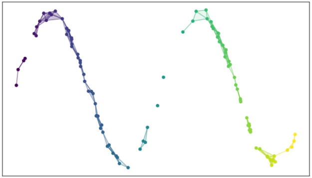


Cech complex　　Vietoris-Rips complex

度量空间中的单纯复形

- **单纯复形与相似图**
  - 样本集的相似图可以看做是单纯复形的简化，只包含了其中的0-simplex和1-simplex
  - Cech复形与数据分布的流形是否具有拓扑一致性？
  - 如何设置适合的邻域半径α，避免相似图的断裂、不连续？





流形的拓扑一致性

- **Cech复形与点集**

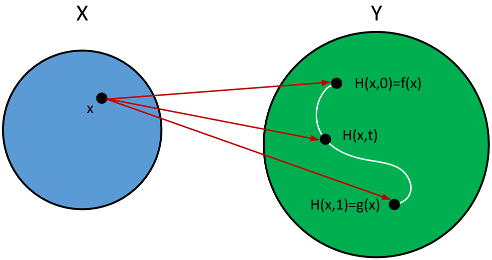
  作为Nerve theorem的推论，如果D是欧氏空间R^d中的点集，则Cech复形$\mathrm{Cech}_\alpha(D)$与$\bigcup_{\mathbf{x}\in D} B(\mathbf{x},\alpha)$是**同伦等价**的

- **同伦(Homotopic)与同胚(Homeomorphic)**
  - 两个拓扑空间是同胚的，可以认为在拓扑意义下是相同的空间，具有共同的拓扑属性
  - 同伦的两个拓扑空间不一定同胚，但也会具有一些相同的拓扑不变量

同伦映射

- **映射的同伦**

  令f, g是由空间X到Y的映射，如果有连续映射$H: X \times I \to Y$，使得$\forall x \in X$，$H(x,0) = f(x), H(x,1) = g(x)$，则称$f$与$g$是同伦映射，记为$f \simeq g: X \to Y$



$H(x,t)$将区间$I = [0,1]$映射为了Y中的一条道路

同伦空间

- **空间的同伦等价**

  令X和Y为拓扑空间，如果有连续映射$f: X \to Y$和$g: Y \to X$，使得：

  $$g \circ f \simeq id_X : X \to X, \qquad f \circ g \simeq id_Y : Y \to Y$$

  则称X和Y是同伦等价的

- **同伦与同胚**
  - 上述的两个复合映射只要求同伦于X和Y空间中的恒等映射id_X, id_Y，而不是相等
  - 两个空间同胚等价必然是同伦等价，因为可以用同胚映射f和f^{-1}作为f和g
  - 同伦不一定同胚，映射f和g不一定是一一映射

同伦/非同胚的空间

- $R^1 \simeq R^2$

  令$f: R^2 \to R^1$为$f(x,y) = x$，$g: R^1 \to R^2$为$g(x) = (x,0)$，易见$f \circ g = id_{R^1}$，下面证明$g \circ f \simeq id_{R^2}$

  构造映射$H: R^2 \times I \to R^1$：

  $$H(x,y,t) = (x,ty) \Longrightarrow H(x,y,0) = (x,0), H(x,y,1) = (x,y)$$

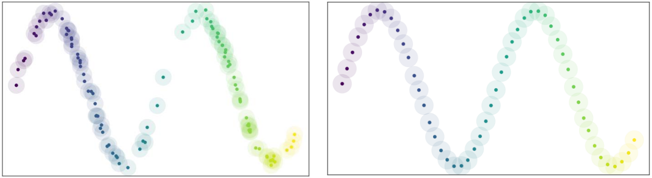
  有：

  $$\begin{aligned} [g \circ f](x,y) &= g(x) = (x,0) = H(x,y,0) \\ id_{R^2}(x,y) &= (x,y) = H(x,y,1) \end{aligned}$$

  H为连续映射，因此有$g \circ f = id_{R^2}$

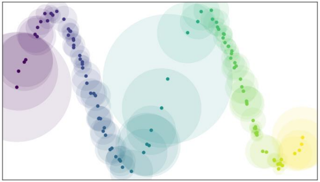
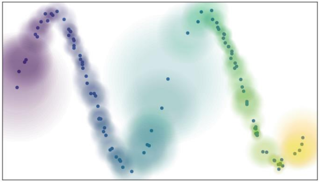
k-近邻相似图

- ϵ-近邻相似图
  - ϵ-近邻相似图是对Cech$_\alpha$复形的近似
  - 当样本集D均匀抽样于流形时，$\bigcup_{\mathbf{x}\in D} B(\mathbf{x},\alpha)$可以很好地描述流形的拓扑特性
  - 当样本集的分布不均匀时，很难选择合适的半径α，容易造成造成相似图的断裂



k-近邻相似图

- k-近邻相似图
  - 流形上每一点的局部黎曼度量是不同的，相同半径的局部球形区域“大小”是不同的
  - 复形是单纯形的集合，以模糊集合代替确定性集合，每条边(1-simplex)上的权重相当于其属于复形的隶属度

流形上的拉普拉斯算子

欧氏空间中的拉普拉斯算子

- **欧氏空间中的梯度算子**
  - 令$f(\mathbf{x})$是R^d空间中的可微函数，梯度算子定义为：

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_d \end{pmatrix}$$

  - 梯度算子∇将函数(标量场)映射为了矢量(矢量场)
  - 梯度$\nabla f(\mathbf{x})$的方向是函数方向导数最大值的方向，长度是该方向上函数的变化率
  - 有时也可以将梯度算子记为一种矢量形式：

$$\nabla = \begin{pmatrix} \partial/\partial x_1 \\ \vdots \\ \partial/\partial x_d \end{pmatrix}$$

欧氏空间中的拉普拉斯算子

- **欧氏空间中的散度**
  - 令$\mathbf{F}$是R^d空间中矢量场，闭合曲面$\mathcal{S}$内部的体积为V
  - 矢量场在闭合曲面的通量(通过闭合曲面的净流出量)：

$$\Phi = \oint_{\mathcal{S}} < \mathbf{F}, d\mathbf{S} >$$

    其中，$d\mathbf{S}$为法矢量面元
  - 矢量场$\mathbf{F}$在点$\mathbf{x}$的散度定义为极限：

$$\mathsf{div}\mathbf{F}(\mathbf{x}) = \lim_{V \to 0} \frac{1}{V} \oint_{\mathcal{S}} < \mathbf{F}, d\mathbf{S} >$$

  - 矢量场的散度可以表示为坐标形式：

$$\mathsf{div}\mathbf{F} =< \nabla, \mathbf{F} >= \sum_{i=1}^{d} \frac{\partial F_i}{\partial x_i}$$

欧氏空间中的拉普拉斯算子

- **欧氏空间中的拉普拉斯算子**
  - 令$f(\mathbf{x})$是定义在开集$\Omega \in R^d$上的二次可微函数
  - 拉普拉斯算子$\Delta : C(\Omega) \rightarrow C(\Omega)$，是一个由标量场到标量场的映射
  - 拉普拉斯算子作用于函数$f(\mathbf{x})$，定义为函数梯度场的散度：

$$\Delta f(\mathbf{x}) = \mathrm{div}(\nabla f(\mathbf{x})) =< \nabla, \nabla f(\mathbf{x}) >= \sum_{i=1}^{d} \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}$$

流形上的拉普拉斯算子

- **流形的切空间**
  - 光滑流形$\mathcal{M}$在点$\mathbf{p} \in \mathcal{M}$的**切向量**$\mathbf{v} : C_{\mathbf{p}}^{\infty} \to R$，为满足下列两个条件的映射
    1. $\mathbf{v}(f + \lambda g) = \mathbf{v}(f) + \lambda \mathbf{v}(g), \quad \forall f, g \in C_{\mathbf{p}}^{\infty}, \forall \lambda \in R$
    2. $\mathbf{v}(fg) = \mathbf{v}(f) g(p) + f(p) \mathbf{v}(g), \quad \forall f, g \in C_{\mathbf{p}}^{\infty}$

    其中，$C_{\mathbf{p}}^{\infty}$表示在$\mathbf{p}$点光滑函数的集合
  - 点$\mathbf{p}$处所有切向量的集合$T_{\mathbf{p}}\mathcal{M}$，称为$\mathcal{M}$在点$\mathbf{p}$的**切空间**

流形上的拉普拉斯算子

- **切空间的基**

  令$(U,\varphi;\{x_1,\cdots,x_m\})$是流形$\mathcal{M}$在$\mathbf{p}$点的一个局部坐标系，沿着第$i$条坐标曲线的切向量记为$\left.\frac{\partial}{\partial x_i}\right|_{\mathbf{p}}$；可以证明，集合

$$\left\{\left.\frac{\partial}{\partial x_i}\right|_{\mathbf{p}}, 1\leq i\leq m\right\}$$

  构成了切空间$T_{\mathbf{p}}\mathcal{M}$的一组基

流形上的拉普拉斯算子

- **光滑切向量场**
  - 流形$\mathcal{M}$上的切向量场$\mathbf{F}:\mathcal{M}\to T\mathcal{M}$，其中
$$T\mathcal{M}=\bigcup_{\mathbf{p}\in\mathcal{M}}T_{\mathbf{p}}\mathcal{M}$$
  - 切向量场$\mathbf{F}$在$\mathcal{M}$的每一点$\mathbf{p}$处指定了一个切向量$\mathbf{F}(\mathbf{p})$
  - 切向量场可以用切空间的基来表示：
$$\mathbf{F}(\mathbf{p})=\sum_{i=1}^{m}F_i(\mathbf{p})\left.\frac{\partial}{\partial x_i}\right|_{\mathbf{p}}$$
  - 如果$F_i(\mathbf{p})$均为光滑函数，则称$\mathbf{F}$为$\mathcal{M}$上的光滑切向量场，$\mathcal{M}$上光滑切向量场的集合记为$\mathfrak{X}(\mathcal{M})$

流形上的拉普拉斯算子

- **切空间中矢量的内积**

  流形$\mathcal{M}$在点p处局部坐标系下，黎曼度量G的元素：

$$g_{ij} = \left\langle \frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right\rangle, \quad i,j = 1, \cdots, m$$

  切空间$T_p\mathcal{M}$中的两个矢量：

$$\mathbf{v} = v_1 \frac{\partial}{\partial x_1} + \cdots + v_m \frac{\partial}{\partial x_m}$$
$$\mathbf{u} = u_1 \frac{\partial}{\partial x_1} + \cdots + u_m \frac{\partial}{\partial x_m}$$

  两个矢量之间的内积：

$$< \mathbf{v}, \mathbf{u} > = \sum_{i=1}^{m} \sum_{j=1}^{m} g_{ij} v_i u_j$$

流形上的拉普拉斯算子

- **流形上的梯度算子**
  - 黎曼流形$(\mathcal{M}, G)$上的梯度算子$\nabla : C^\infty(\mathcal{M}) \to \mathfrak{X}(\mathcal{M})$，将流形上的标量场映射为切向量场
  - 令$f \in C^\infty(\mathcal{M})$是流形$\mathcal{M}$上的可微函数，在局部坐标系下，函数的梯度可以表示为：

$$\begin{aligned}\nabla f &= \left(\sum_{i=1}^{m} g^{i1} \frac{\partial f}{\partial x_i}\right) \frac{\partial}{\partial x_1} + \cdots + \left(\sum_{i=1}^{m} g^{im} \frac{\partial f}{\partial x_i}\right) \frac{\partial}{\partial x_m} \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{m} g^{ij} \frac{\partial f}{\partial x_i}\right) \frac{\partial}{\partial x_j}\end{aligned}$$

其中，g^{ij}是黎曼度量G对应矩阵的逆矩阵G^{-1}的元素

流形上的拉普拉斯算子

- **流形上的散度算子**
  - 黎曼流形$(\mathcal{M}, G)$上的散度算子div : $\mathfrak{X}(\mathcal{M}) \to C^{\infty}(\mathcal{M})$，将流形上的切向量场映射为标量场
  - 令$\mathbf{F} \in \mathfrak{X}(\mathcal{M})$ 是流形$\mathcal{M}$上的切向量场，在局部坐标系下，切向量场的散度可以表示为：

$$\mathrm{div}\mathbf{F} = < \nabla, \mathbf{F} > = \frac{1}{\sqrt{|G|}} \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\sqrt{|G|} F_i \right)$$

其中，$|G|$是黎曼度量对应矩阵的行列式值

流形上的拉普拉斯算子

- **拉普拉斯-贝尔特拉米算子**
  - 黎曼流形$(\mathcal{M}, G)$上的算子$\Delta : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$，将流形上的标量场映射为标量场，称为Laplace－Beltrami算子
  - 令$f \in C^\infty(\mathcal{M})$是流形$\mathcal{M}$上的光滑函数，在局部坐标系下：

$$\Delta f =< \nabla, \nabla f >= \frac{1}{\sqrt{|G|}} \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\sqrt{|G|} \sum_{j=1}^{m} g^{ij} \frac{\partial f}{\partial x_j} \right)$$

Stokes定理

- **Stokes' Theorem**

  负散度算子$-\mathrm{div}$与梯度算子∇为伴随算子(formally adjoint operators)，满足：

$$\int_{\mathcal{M}} \langle \mathbf{F}, \nabla f \rangle \, d\mathbf{x} = -\int_{\mathcal{M}} \mathrm{div}\mathbf{F} \cdot f d\mathbf{x}$$

  当切向量场$\mathbf{F} = \nabla f$时，有：

$$\int_{\mathcal{M}} \|\nabla f\|^2 d\mathbf{x} = \int_{\mathcal{M}} \langle \nabla f, \nabla f \rangle \, d\mathbf{x} = -\int_{\mathcal{M}} \mathrm{div}(\nabla f) \cdot f d\mathbf{x} = -\int_{\mathcal{M}} \Delta f \cdot f d\mathbf{x}$$

流形的低维嵌入

- **流形的低维映射**

  令$\mathcal{M}$是嵌入在R^d空间的m维黎曼流形，希望寻找流形到1维欧氏空间的映射$f:\mathcal{M}\to R$，使得流形上临近的点$\mathbf{x},\mathbf{z}$，映射之后的距离尽可能地小

  上述目标同谱嵌入的框架是一致的，差别在于这里讨论的是连续的低维流形嵌入，而谱嵌入框架面对的是流形上的一个抽样数据集

流形的低维嵌入

- **映射之后的距离**

  令$l = \mathsf{dist}_{\mathcal{M}}(\mathbf{x}, \mathbf{z})$为两点在流形上的测地距离，$c(t)$为参数化的测地曲线，则有：

$$f(\mathbf{z}) = f(\mathbf{x}) + \int_0^l df(c'(t))dt$$

  其中，$df(c'(t))$是映射f沿着测地曲线切线方向的增量，积分由测地线的一端到另一端累计函数的增量

  利用Schwartz不等式和Taylor近似可以得到，映射之后两点之间的距离：

$$|f(\mathbf{z}) - f(\mathbf{x})| \leq l\|\nabla f(\mathbf{x})\| + o(l)$$

流形的低维嵌入

- **优化低维映射**

  $l = \mathrm{dist}_{\mathcal{M}}(\mathbf{x}, \mathbf{z})$是两点在流形上的测地距离，与映射$f$无关，因此希望映射之后两点的距离尽量小，需要最小化整个流形上每个点梯度的范数

  为了避免得到$f(\mathbf{x}) = C$的无意义解，需要约束函数$f(\mathbf{x})$的范数：

$$\arg\min_{\|f\|^2=1} \int_{\mathcal{M}} \|\nabla f(\mathbf{x})\|^2 d\mathbf{x}$$

  根据Stokes定理，得到等价的优化问题：

$$\arg\max_{\|f\|^2=1} \int_{\mathcal{M}} \Delta f(\mathbf{x}) \cdot f(\mathbf{x}) d\mathbf{x}$$

流形的低维嵌入

- **嵌入函数与拉普拉斯算子的谱**

  构造Lagrange函数：

$$L\left(f(\mathbf{x})\right) = \int_{\mathcal{M}} \Delta f(\mathbf{x}) \cdot f(\mathbf{x}) d\mathbf{x} - \lambda(\|f(\mathbf{x})\|^2 - 1)$$

  对$f(\mathbf{x})$求导，计算极值：

$$\Delta f(\mathbf{x}) = \lambda f(\mathbf{x})$$

  因此，嵌入函数$f(\mathbf{x})$为拉普拉斯算子Δ的特征函数

流形的低维嵌入

- **拉普拉斯算子与拉普拉斯矩阵**

  如果有流形$\mathcal{M}$上的有限抽样集$\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$，可以构造近邻相似图(1-simplical complex)近似流形的拓扑结构；函数$f(\mathbf{x})$映射之后的函数值为$\{f_1, \cdots, f_n\}$

  沿着近邻$\mathbf{x}_i$到$\mathbf{x}_j$测地线c_{ij}方向上函数的变化率：

$$\nabla_{c_{ij}} f(\mathbf{x}_i) \approx \frac{f_j - f_i}{\|\mathbf{x}_i - \mathbf{x}_j\|_{\mathcal{M}}}$$

  如果令$W_{ij} \approx \frac{1}{\|\mathbf{x}_i - \mathbf{x}_j\|^2_{\mathcal{M}}}$，函数$f(\mathbf{x})$在$\mathbf{x}_i$点梯度长度的平方，可以用各个方向上函数变化率平方之和来近似：

$$\|\nabla f(\mathbf{x}_i)\|^2 \approx \sum_{\mathbf{x}_j \in N(\mathbf{x}_i)} (f_i - f_j)^2 W_{ij}$$

流形的低维嵌入

- **拉普拉斯算子与拉普拉斯矩阵**

  用抽样集上梯度长度平方之和近似在整个流形上的积分：

$$\int_{\mathcal{M}} \|\nabla f(\mathbf{x})\|^2 d\mathbf{x} \approx \sum_{i,j} (f_i - f_j)^2 W_{ij} = \mathbf{f}^t L \mathbf{f}$$

  比较Stokes公式：

$$\int_{\mathcal{M}} \|\nabla f\|^2 d\mathbf{x} = \int_{\mathcal{M}} f \cdot \Delta f d\mathbf{x}$$

  可以看出，相似图的拉普拉斯矩阵L是流形$\mathcal{M}$上拉普拉斯-贝尔特拉米算子Δ的近似