Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions

$(x_1, y_1) \ldots (x_l, y_l)$ 打标签数据

$x_{l+1}, \ldots x_{l+u}$ 未打标签的数据

$x \in R^m$

$$w_{ij} = \exp\left(-\sum_{d=1}^{m} \frac{(x_{id} - x_{jd})^2}{\sigma_d^2}\right) \qquad (1)$$

刻画xi，xj之间的相似度

因为如果xi==xj,那么wij就是1.
否则就是小于1.

$$E(f) = \frac{1}{2} \sum_{i,j} w_{ij} \left(f(i) - f(j)\right)^2 \qquad (2)$$

f是分类函数输入xi输出yi,并且我们需要保证两个点越近f越近.

能量函数的定义：
将待聚类的事物看成一个系统，事物之间的相异程度看成系统元素间的能量，当能量达到一定程度时，事物就形成一个新的类，表示系统需要重新分类。聚类过程中要求每个事物属于一个类，每个簇中不存在能量大于阈值的系统，不同的簇中不存在能量小于阈值的的系统。

为什么E里面需要wij和f2个度量的乘积.因为我们需要衡量的是一个多元素的聚类效果问题.
所以我们需要f效果差的时候E(f)很大. 所以需要wij来让xi,xj远的时候差， fi-fj让聚类效果差的时候E大.

图上 Fourier

G: 无向图

$L = D - A$ 度矩阵 - 邻接矩阵

$\because L$ 是对称阵

$\therefore L$ 存在酉矩阵使得他对角化.

$$L = U \Lambda U^T$$

图象上: $\hat{f}(t) = \int f(x) \underbrace{\exp^{-2\pi i x t}} dx$

$\exp^{-2\pi i x t}$ 是 t域上的一个 Δ 的特征函数. 因为他对t求二阶导数是他自己的常熟北. 这里面x看做常数.

$$\Delta \exp^{-2\pi i x t} = \frac{\partial^2}{\partial t^2} \exp^{-2\pi i x t} = -4\pi^2 x^2 \exp^{-2\pi i x t}$$

$\therefore$ 图上 $\hat{f}(t) = \sum f(n) U_t(n)$

n 取遍所有的节点.

记 $\hat{f} = U^T f$ 乘以U就是fourier变换

$$f * g = U\left(U^T f \odot U^T g\right)$$ 傅里叶恒等式

$$= U g_\theta U^T f$$

$$g_\theta \triangleq U^T g$$ 作为学习系数

$$L = U \Lambda U^T = I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$$

D是对角矩阵.所以他的元素存在 $D^{-\frac{1}{2}}$ 对角线元素分别-0.5次幂即可

证明等式：

$$L = D - A$$

$D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$ 特征值在0,1之间.

证: $D^{-\frac{1}{2}} A D^{-\frac{1}{2}} x = \lambda x$

$A D^{-\frac{1}{2}} x = \lambda D^{\frac{1}{2}} x$

$D^{-\frac{1}{2}} x \triangleq y$; $A y = \lambda D y$

根据D的定义,D里面对角线是A的该列的sum, 所以lamda显然0,1之间.

一个laplace算子的推导过程:
https://zhuanlan.zhihu.com/p/81370490

$n! \times (n) \times (n-1)$

$n = 2$ ✓

$n = 4$ AB CD

2

① 6

② 2

③ 12

AB CD

$2 \times 2 \times 2$

④ $5!/2! = 3 \times 4 \times 5$

$= 60$

$1000 = 1000$

$= \cancel{499 + 500}$

$= \cancel{333\ 333\ 334}$

$= 5$个

$= 13$

$=$

$2 \,\underline{|\, 1000} =$

$2 \,\underline{|\, 500} = 5 \times 5 \times 4$

$= \times 5$

1400 = 2 x 5

1 5 5^2 5^3

2 3 5

2042

2^p 3^p 5

x 1~x 有多少

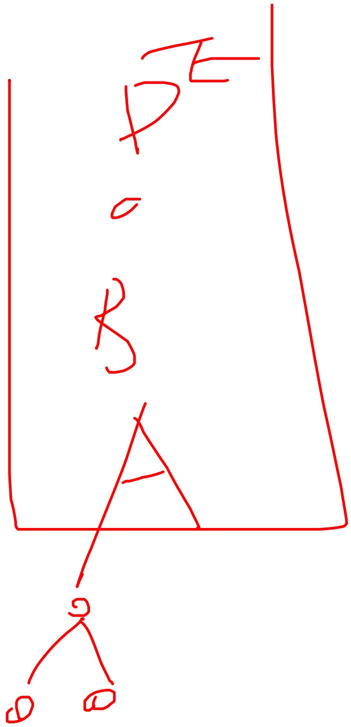
2倍 3倍 5倍 - 6倍 - 15倍

2040 有

$1020+680+408$
$-340-136-204+68$
$=1496$ 2040

2042 2043 2044

2041



360

$1+2+4$

$2^0+2^1+\cdots+2^8=360$

$2^{n+1}-1=360$

$2 \,|\, 361 \qquad n+1=9$

$n=8$

180.5

$2^3=8$

$2^{6} = 64$

$2^{7} = 128$

$2^{8} = 256$

$2^{9} = 4\ldots$

100 100 100 100

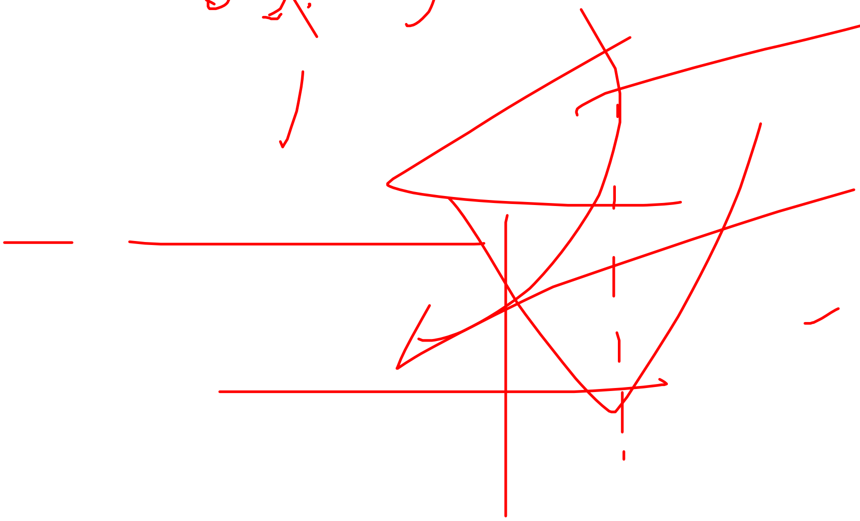
1 1 1 1

$100 \rightarrow 100, 100, 100$

5 6 7 8 Hoffman

$$= \cancel{\sum_{j=1}^{T}}[(\sum_{i\in I_j} g_i)w_j + \frac{1}{2}(\sum_{i\in I_j} h_i + \lambda)w_j^2] + \gamma T$$

$$\frac{1}{2}(\sum_{i\in I_j} h_i + \lambda) w_j^2 + \sum_{i\in I_j} g_i$$



$-\frac{b}{2a}$

$ax^2 + bx + c = 0$