

# 基于辅助空间与主空间合作的半监督聚类方法

罗晓清, 王士同, 徐红林

LUO Xiao-qing, WANG Shi-tong, XU Hong-lin

江南大学 信息工程学院, 江苏 无锡 214122

School of Information Technology, Southern Yangtze University, Wuxi, Jiangsu 214122, China

E-mail: zdlxq@sohu.com

LUO Xiao-qing, WANG Shi-tong, XU Hong-lin. Novel semi-supervised clustering: collaborating primary space with auxiliary space. *Computer Engineering and Applications*, 2007, 43(23): 177-180.

**Abstract:** In general, unsupervised clustering is only based on primary space. If auxiliary space is considered to incorporate with primary space, unsupervised clustering will become semi-supervised clustering. In this paper, auxiliary space combined with primary space is introduced to the design of the corresponding cost function. And accordingly a novel semi-supervised clustering algorithm APMSC based on both auxiliary space and primary space is proposed. This algorithm realizes the efficient clustering by minimizing the corresponding cost function iteratively. Our experimental results demonstrate its validity and superiority.

**Key words:** auxiliary space; primary space; cost function; semi-supervised clustering

**摘要:** 无监督聚类仅仅基于主空间, 当辅助空间被引入聚类过程时, 无监督聚类成为半监督聚类。代价函数的设计既考虑到主空间又考虑到辅助空间, 从而一个新颖的基于辅助空间与主空间合作的半监督聚类方法 APMSC 被提出。该算法通过迭代优化, 使得相应的代价函数最小化, 最终得到有效的聚类结果。通过实验证实了 APMSC 的有效性和优越性。

**关键词:** 辅助空间; 主空间; 代价函数; 半监督聚类

文章编号: 1002-8331(2007)23-0177-04 文献标识码: A 中图分类号: TP181

## 1 引言

聚类分析是对群体及成员进行分类的递归过程。一个簇是一组数据对象的集合, 在同一簇中的对象彼此类似, 而不同簇中的对象彼此相异。聚类分析有很广泛的应用, 包括市场或客户分析、模式识别、生物学研究、Web 文档分类等。

Janne SinKKonen 和 Samuel Kaski 提出了一种基于辅助空间条件分布的半监督聚类方法 AMSC<sup>[1-3]</sup>。该算法中主空间的信息被投影到辅助空间, 通过辅助空间完成迭代优化过程。AMSC 中数据符合 VMF 分布, 每个样本事先不知道分类。经研究表明同类样本具有相似的相关信息, 因此作者引入辅助空间对数据进行监督和限制, 提高聚类的准确性。

虽然, 通过 Janne SinKKonen 和 Samuel Kaski 的论证, 知道 AMSC 辅助空间的引入较好地改善了聚类效果, 但是主空间的变化对聚类过程的影响被削弱, 不能很好地体现在迭代优化中。因而目标函数往往收敛于鞍点, 不能有效地处理局部极值问题, 聚类效果不稳定, 如何避免或改进这些问题是本文的目标所在。本文的工作将基于辅助空间与主空间的两个相似函数相结合, 提出 APMSC 聚类算法。该算法引入两个距离函数, 重新构造代价函数, 通过迭代优化公式的改进, 使得聚类的过程受两个空间和两个相似函数制约。并通过实验证明了该算

法聚类性能优于 AMSC。

本文由以下几个部分组成, 第 2 章介绍数据的 VMF 分布, 第 3 章描述了 AMSC 算法, 第 4 章提出了 APMSC 算法, 以及算法推导。第 5 章给出聚类评价标准和实验结果。最后对本文进行了总结和展望。

## 2 THE VON MISES-FISHER (VMF) 分布

THE VON MISES-FISHER (VMF) 分布<sup>[4]</sup>是最常见的方向性数据分布。它类似于在超球面上多维的高斯分布。 $n$  维 VMF 分布的概率密度函数如下:

$$VMF(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{Z_n(k)} \exp k \frac{\mathbf{x}^T \boldsymbol{\theta}}{\|\boldsymbol{\theta}\|} \quad (1)$$

该函数中, 参数向量  $\boldsymbol{\theta}$  代表平均方向向量是位置参数, VMF 分布中心位置由它确定。 $k$  为尺度参数, 分布的宽度由参数  $k$  确定。

在本文的聚类算法中, 使用 VMF 核函数:

$$f(\mathbf{x}; \boldsymbol{\theta}_i) = \frac{k \mathbf{x}^T \boldsymbol{\theta}_i}{\|\boldsymbol{\theta}_i\|} \quad (2)$$

设  $k$  为常量, 归一化  $\boldsymbol{\theta}_i$ ,

$$\frac{\partial f(\mathbf{x}; \boldsymbol{\theta}_i)}{\partial \boldsymbol{\theta}_i} = k (\mathbf{x} - \mathbf{x}^T \boldsymbol{\theta}_i \boldsymbol{\theta}_i) \quad (3)$$

**基金项目:** 教育部新世纪优秀人才计划; 教育部科学研究重点项目; 模式识别国家重点实验室开放基金。

**作者简介:** 罗晓清(1980-), 女, 硕士研究生, 主要研究方向: 模式识别, 生物信息学; 王士同(1964-), 男, 教授, 博士生导师, 主要研究方向: 模式识别、人工智能、生物信息学; 徐红林(1981-), 女, 博士研究生, 主要研究方向: 模式识别、生物信息学。

### 3 AMSC 算法

向量量化(VQ)<sup>[5]</sup>或者 K-means 聚类,都是分类的方法。在 VQ 中,目标是使数据和原型或码书(向量  $m_j$ )之间的平均失真  $E$  最小化。定义如下:

$$E = \sum_j \int y_j(x) D(x, m_j) p(x) dx \quad (4)$$

在 AMSC 算法<sup>[1-3]</sup>中,距离度量采用的是 Kullback-Leibler<sup>[6]</sup>距离,定义两个离散分布的事件概率为  $\{p_i\}$  和  $\{\varphi_i\}$ 。K-L 距离的形式为:

$$D_{KL}(p, \varphi) \equiv \sum_i p_i \text{lb} \left( \frac{p_i}{\varphi_i} \right) \quad (5)$$

在 AMSC 中,  $D$  是指  $p(c|x)$  和模型分布之间的不同。  $P_i = p(c_i|x)$ , 它表示样本  $x$  出现  $c_i$  这个信息的概率。第二个分布  $\varphi_i$ ,  $\varphi_i = p(c_i|v_j)$ , 表示聚类  $v_j$  出现  $c_i$  这个信息的概率。

当把 Kullback-Leibler 失真度量方式插入到 VQ 代价函数(4)中,  $y_j(x)$  由参数  $\theta_j$  来表示时,代价函数可表示为以下形式:

$$E_{KL} = - \sum_{i,j} \int [y_j(x; \theta_j) \text{lb} \varphi_{ij}] p(c_i, x) dx + const \quad (6)$$

这里常量  $const$  是独立于参数的。随机变量  $v_j$  是聚类的标识,随机变量  $c_i$  是多项式辅助随机变量。假设辅助信息有  $m$  个取值,表示为  $(c_1, c_2, \dots, c_m)$ ,  $i=1, \dots, m$ 。隶属度函数  $y_j(x; \theta_j)$  可以看作是条件概率  $p(v_j|x) = y_j(x)$ , 表示  $x$  属于类  $v_j$  的概率。因为数据满足 VMF 分布,隶属度函数  $y_j(x; \theta_j)$  进行指数归一化后形式为:

$$y_j(x; \theta_j) = \frac{\exp f(x; \theta_j)}{\sum_i \exp f(x; \theta_j)} \quad (7)$$

其中  $f(x; \theta_j)$  是 VMF 概率分布核函数,公式(6)可以表示为:

$$E_{KL} = - \sum_{i,j} p(c_i, v_j) \text{lb} \frac{p(c_i, v_j)}{p(c_i)p(v_j)} + const = -I(C; V) + const \quad (8)$$

$I(C; V)$  表示随机变量  $c$  和  $v$  的互信息<sup>[6,7]</sup>。从中可以看出,最小平均失真等于辅助变量  $c$  和聚类  $v$  之间的最大互信息。

### 4 APMSC 算法

AMSC 算法把辅助空间引入无监督聚类中,形成新的半监督聚类算法,经 Janne Sinkkonen 和 Samuel Kaski 验证,它提高了收敛速度,改善了聚类性能。但是 AMSC 只是把主空间投影到辅助空间,通过辅助空间完成聚类,这种算法较多地考虑辅助空间的变化,而削弱了主空间的影响,在迭代过程中算法不够稳定,往往收敛于局部最小,导致早熟现象。经研究发现主空间和辅助空间共同影响着聚类过程,应该把两者融合到算法中,从而得到更为有效的聚类结果。

辅助空间和主空间相结合,辅助空间采取 Kullback-Leibler 距离,主空间采取欧氏距离。公式(6)变为:

$$E_{KL} = - \sum_{i,j} \int [y_j(x; \theta_j) \text{lb} \varphi_{ij}] p(c_i, x) dx + \sum_j \int y_j(x; \theta_j) \|x - \theta_j\|^2 p(x) dx + const \quad (9)$$

下面,针对公式(9),分别对  $\theta_l$  和  $\gamma_{lm}$  求导:

$$\text{令 } A = - \sum_{i,j} \int [y_j(x; \theta_j) \text{lb} \varphi_{ij}] p(c_i, x) dx, \text{ 同时令 } B = \sum_j \int y_j(x;$$

$\theta_j) \|x - \theta_j\|^2 p(x) dx$ , 其中:

$$\frac{\partial A}{\partial \theta_l} = \sum_{i,j} \int \frac{\partial f(x; \theta_l)}{\partial \theta_l} \text{lb} \frac{\varphi_{ij}}{\varphi_i} p(c_i, v_j, v_l, x) dx \quad (10)$$

$$\frac{\partial y_j(x; \theta_j)}{\partial \theta_l} = \delta_{jl} y_j(x; \theta_j) \frac{\partial f(x; \theta_j)}{\partial \theta_l} - y_j(x; \theta_j) y_l(x; \theta_l) \frac{\partial f_l(x; \theta_l)}{\partial \theta_l} \quad (11)$$

$\delta_{jl} = 1$ , 当  $j=l$  时;  $\delta_{jl} = 0$ , 当  $j \neq l$  时。

$$\text{令 } K = \sum_j y_j(x; \theta_j) \|x - \theta_j\|^2, \frac{\partial B}{\partial \theta_l} = \int \frac{\partial K}{\partial \theta_l} p(x) dx$$

$$\frac{\partial K}{\partial \theta_l} = \sum_j \frac{\partial [y_j(x; \theta_j) \|x - \theta_j\|^2]}{\partial \theta_l} = \sum_j \left( \frac{\partial y_j(x; \theta_j)}{\partial \theta_l} \|x - \theta_j\|^2 - 2y_j(x; \theta_j)(x - \theta_l) \right)$$

把公式(11)代入上式中,  $y_j(x)$  是隶属度函数,  $\sum_j y_j(x; \theta_j) = 1$ , 可以得到:

$$\frac{\partial K}{\partial \theta_l} = \sum_j y_j(x; \theta_j) y_l(x; \theta_l) \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_l\|^2 - \sum_j y_j(x; \theta_j) y_l(x; \theta_l) \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_j\|^2 - 2 \sum_j y_j(x; \theta_j) y_l(x; \theta_l) (x - \theta_l) \quad (12)$$

$$\therefore \frac{\partial B}{\partial \theta_l} = \sum_j \int [y_j(x; \theta_j) y_l(x; \theta_l) \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_l\|^2 - y_j(x; \theta_j) y_l(x; \theta_l) \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_j\|^2 - 2y_j(x; \theta_j) y_l(x; \theta_l) (x - \theta_l)] \sum_i p(c_i|x) p(x) dx \quad (13)$$

因为关于  $x, v$  和  $c$  条件独立, 所以

$$y_j(x; \theta_j) y_l(x; \theta_l) p(c_i, x) = p(c_i, v_j, v_l, x) \quad (14)$$

$$\therefore \frac{\partial B}{\partial \theta_l} = \sum_{i,j} \int \left[ \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_l\|^2 - \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_j\|^2 - 2(x - \theta_l) \right] p(c_i, v_j, v_l, x) dx \quad (15)$$

$$\frac{\partial E_{KL}}{\partial \theta_l} = \sum_{i,j} \int \left[ \frac{\partial f(x; \theta_l)}{\partial \theta_l} \text{lb} \frac{\varphi_{ij}}{\varphi_i} + \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_l\|^2 - \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_j\|^2 - 2(x - \theta_l) \right] p(c_i, v_j, v_l, x) dx \quad (16)$$

现在对  $\gamma_{lm}$  进行求导:

原型  $\varphi_j$  是多项式概率分布, 为了使它满足  $0 \leq \varphi_i \leq 1$  和

$\sum \varphi_i = 1$ 。对原型参数化, 得到:

$$\text{lb} \varphi_{ij} \equiv \gamma_{ij} - \text{lb} \sum_m \exp(\gamma_{jm}), \frac{\partial B}{\partial \gamma_{lm}} = 0 \quad (17)$$

$$\frac{\partial E_{KL}}{\partial \gamma_{lm}} = \sum_i \int (\varphi_{lm} - \delta_{mi}) p(c_i, v_l, x) dx \quad (18)$$

$\delta_{mi} = 1$ , 当  $m=i$  时;  $\delta_{mi} = 0$ , 当  $m \neq i$  时。

代价函数采用随机逼近方法<sup>[8]</sup>最小化。  $t$  表示迭代次数,  $\alpha(t)$  表示学习率, 是一个递减到 0 的函数。  $\alpha(t)$  满足以下条件:

$$\sum_t \alpha(t) = \infty \text{ 和 } \sum_t \alpha(t)^2 < \infty$$

下面的步骤按照  $t=0, 1, \dots$  和  $\alpha(t)$  递减至 0 来循环。

(1) 在随机逼近的步骤  $t$  中, 随机抽取数据样本  $x_t$ , 然后在  $x_t$  对应的辅助变量值中随机抽取  $c_t$ , 形成数据集  $(x_t, c_t)$ 。

(2) 针对抽取的样本  $(x_t, c_t)$ , 计算  $x_t$  相对于每个子类的隶属度  $\{y_k(x(t))\}_k$ , 以隶属度为  $K$  个子类出现的概率, 随机地按概率  $y_k$  抽取其中两个子类, 记为  $j$  和  $l$ 。

(3) 根据下面两个公式, 调整参数  $\theta_l$  和  $\gamma_{lm}, m=1, \dots, N_c$ 。

$$\theta_l(t+1) = \theta_l(t) - \alpha(t) \left[ \frac{\partial f(x; \theta_l(t))}{\partial \theta_l(t)} \frac{\varphi_{ji}}{\varphi_{li}} + \frac{\partial f(x; \theta_l)}{\partial \theta_l} \|x - \theta_l\|^2 - \frac{\partial f(x; \theta_j)}{\partial \theta_j} \|x - \theta_j\|^2 - 2(x - \theta_l) \right]$$

$$\gamma_{lm}(t+1) = \gamma_{lm}(t) - \alpha(t) (\varphi_{lm} - \delta_{mi}) \quad (19)$$

$N_c$  是指辅助空间随机变量  $c$  的个数,  $j$  和  $l$  是聚类的标识, 由于公式 (19) 中  $j$  和  $l$  具有对称性, 可以在同一次迭代  $t$  中, 通过交换  $j$  和  $l$ , 调整两次。

迭代结束的条件为下列条件之一:

- (1) 超过迭代的最大次数 (迭代的最大次数由用户指定);
- (2) 经过多次迭代之后, 中心向量的变化很小;
- (3) 经过多次迭代之后, 聚类中的数据向量不再变化。

## 5 相关实验与分析

### 5.1 聚类评价标准

采用的是 VMF 数据, 所以可以使用文献 [4, 9] 中的  $H_{AVE}$ 、 $S_{AVE}$  来评价聚类算法的聚类质量。设  $\theta_j$  为归一化的类  $V_j$  的中心, 则  $H_{AVE}$ 、 $S_{AVE}$  定义如下:

$$H_{AVE} = \frac{1}{N} \sum_{i=1}^k \sum_{x \in V_i} x^T \theta_i$$

$$S_{AVE} = \frac{1}{\sum_{i \neq j} |V_i| |V_j|} \sum_{i \neq j} |V_i| |V_j| \theta_i^T \theta_j \quad (20)$$

说明:  $|V_j|$  表示类  $V_j$  中包括的样本数,  $N$  为总的样本个数,  $K$  为类别数。  $H_{AVE}$  用来度量方向性数据类内的样本之间的致密度, 越高说明类内的样本之间的致密度较好。  $S_{AVE}$  用来度量类

间的分离度, 越低说明类间分离较好。一般来说, 在进行聚类时追求高  $H_{AVE}$  值和低  $S_{AVE}$  值。

同时还使用辅助变量和聚类之间的互信息 MI 作为评价标准, 由公式 (8) 可以得出互信息越大说明平均失真越小。

### 5.2 实验

#### 5.2.1 模拟数据 1

(1) 设计 300 个 2 维样本点, 令  $k=30$ , 分成 3 类, 可生成图 1 的 VMF 数据。

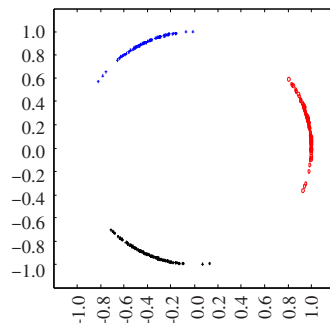


图 1 VMF 数据

表 1 300 个 3 类模拟数据 10 次取平均值聚类评价

| 算法    | 指标      |           |           |
|-------|---------|-----------|-----------|
|       | MI      | $H_{AVE}$ | $S_{AVE}$ |
| AMSC  | 0.402 5 | 0.788 8   | -0.226 1  |
| APMSC | 0.429 5 | 0.981 9   | -0.462 4  |

(2) 模拟辅助空间, 设辅助空间有 5 个功能, 第一类具备 1、2、3 号功能, 第二类具备 1、4 号功能, 第三类具备 3、5 号功能。实验结果见图 2~图 4。

#### 5.2.2 模拟数据 2

(1) 设计 300 个 4 维样本点, 令  $k=30$ , 分成 3 类。

(2) 模拟辅助空间, 设辅助空间有 4 个功能, 第一类具备 1、2 号功能, 第二类具备 3、4 号功能, 第三类具备 2、3 号功能。实验结果见图 5~图 7。

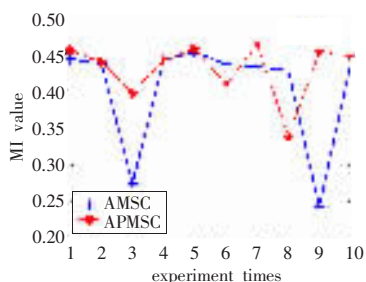


图 2 互信息值 (10 次实验)

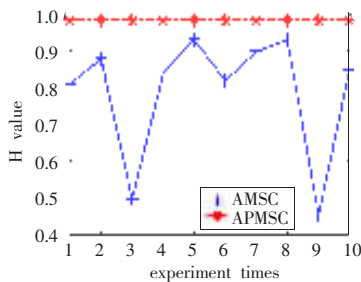


图 3 H 值 (10 次实验)

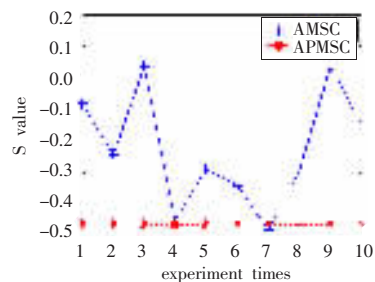


图 4 S 值 (10 次实验)

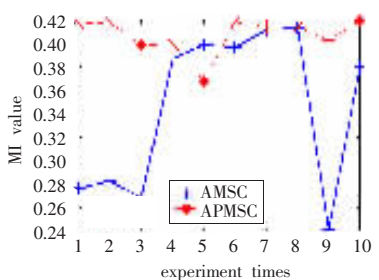


图 5 互信息值 (10 次实验)

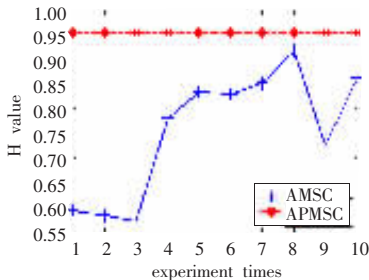


图 6 H 值 (10 次实验)

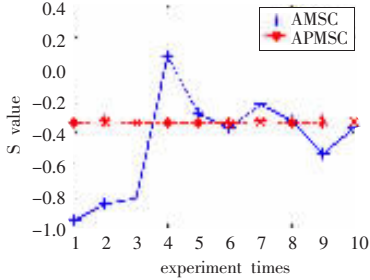


图 7 S 值 (10 次实验)

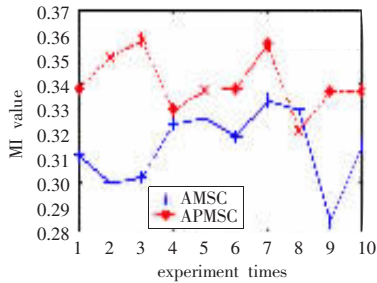


图8 互信息值(10次实验)

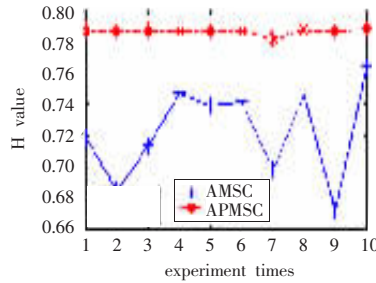


图9 值(10次实验)

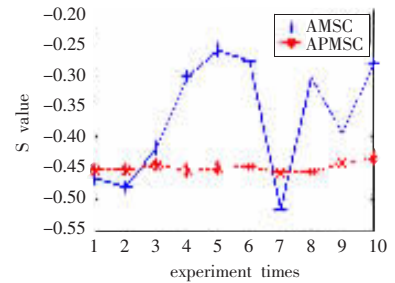


图10 值(10次实验)

表2 300个3类模拟数据10次取平均值聚类评价

| 算法    | 指标      |           |           |
|-------|---------|-----------|-----------|
|       | MI      | $H_{AVE}$ | $S_{AVE}$ |
| AMSC  | 0.345 9 | 0.754 6   | -0.465 1  |
| APMSC | 0.407 1 | 0.954 1   | -0.337 2  |

### 5.2.3 yeast基因数据<sup>[10]</sup>

yeast 基因数据可从 <http://rana.Stanford.edu/clustering/> 中获得, 根据选取的基因通过 the Munich Information Center for Protein Sequences Yeast Genome Database (MYGD) 查找对应的基因功能作为辅助空间 (<http://mips.gsf.de/proj/funecatDB>)。

(1) 选取 384 个 79 维 yeast 基因, 令  $k=148$ , 分成 5 类。

(2) 以每个基因对应的基因功能作为辅助空间。实验结果见图 8~图 10。

表3 384个5类模拟数据10次取平均值聚类评价

| 算法    | 指标      |           |           |
|-------|---------|-----------|-----------|
|       | MI      | $H_{AVE}$ | $S_{AVE}$ |
| AMSC  | 0.314 2 | 0.722 5   | -0.372 6  |
| APMSC | 0.340 3 | 0.786 4   | -0.449 5  |

## 5.3 实验分析

实验 1, 2 是模拟 VMF 数据, 所以  $k$  值是已知的。模拟数据 1 是二维数据因此能够用图形显示出来。AMSC 和 APMSC 都运行 10 次, 最终结果取平均值。从图 2 中可以看出, AMSC 中有 2 个点陷入局部最小, APMSC 中的 MI 值较平稳受初始点影响不大, 图 3, 4 中  $H$  值整体较优,  $S$  值在 10 次运行中基本都小于 AMSC 中的  $S$  值,  $H$  值和  $S$  值非常平稳。同样, 数据 2 也是模拟数据, 只是设为多维。从图 5 中可以看出 AMSC 收敛不太稳定, 有 4 个 MI 值取值很小, APMSC 中 MI 值较稳定, 波动不大, 大部分值大于 AMSC。从图 6 图 7 中可以看出, APMSC 的  $H$  值 10 次实验全大于 AMSC,  $H$  和  $S$  值非常稳定。虽然 AMSC 中  $S$  值有 3 次达到较小值, 但是通过研究发现, 这 3 次对应的 MI,  $H$  值均较小, 所以综合来看这 3 个点并不是  $S$  的正常取值。

实验 3 采用基因数据, 基因数据具有高维量大的特点, 符合 VMF 分布。首先通过实验选取合适的  $k$  值。选取了恰当参数后, 得到如图 8, 9, 10 实验结果。可以看出在 APMSC 中 MI 有 9 个点大于 AMSC, 并且波动范围很小, AMSC 中很明显有些点取值很小, 收敛过早, 陷入局部最小。APMSC 中的  $H$  值大于 AMSC,  $S$  值整体比 AMSC 的小, 同样也比较平稳。

通过上述实验可以看出, 使用 APMSC 算法, 寻找合适的  $k$

值和下降函数, 辅助空间和主空间合作参与到聚类的学习过程中, 样本不管是二维还是多维, 聚类效果从量化误差和类内、类间距离来说都明显优于 AMSC, 并且不易陷入局部最优, 取值较平稳, 受初始点影响较小, 提高了聚类的有效性。

## 6 结语

本文提出了一种基于辅助空间与主空间合作的半监督聚类方法。采用随机逼近方法优化代价函数, 完成聚类过程, 从量化误差、类内、类间距离三个方面证明聚类效果得到了改善。在后续工作中, 将着眼考虑如何快速的选取合适的尺度参数  $K$ 。(收稿日期: 2006 年 11 月)

## 参考文献:

- [1] Sinkkonen J, Kaski S. Clustering based on conditional distributions in an auxiliary space[J]. Neural Computation, 2001(14): 217-239.
- [2] Sinkkonen J, Kaski S. Clustering by similarity in an auxiliary space[C]// Proceedings of IDEAL 2000, Second International Conference on Intelligent Data Engineering and Automated Learning. [S.l.]: Springer, 2000.
- [3] Peltonen J. Data exploration with learning metrics[D]. Helsinki University of Technology, Dissertations in Computer and Information Science, Report D7, Espoo, Finland, 2004.
- [4] Banerjee A, Dhillon I S, Ghosh J, et al. Clustering on the unit hypersphere using von Mises-Fisher distributions[J]. Journal of Machine Learning Research, 2005(6): 1345-1382.
- [5] Theodoridis S, Koutroumbas K. 模式识别[M]. 李晶皎, 朱志良, 王爱侠, 等译. 2 版. 北京: 电子工业出版社, 2004.
- [6] 田宝玉. 工程信息论[M]. 北京: 北京邮电大学出版社, 2004.
- [7] Duda R O, Hart P E, Stork D. 模式分类[M]. 李宏东, 姚天翔, 译. 2 版. 北京: 机械工业出版社, 中信出版社, 2003.
- [8] Kaski S. Convergence of a stochastic semisupervised clustering algorithm, Technical Report A62[R]. Helsinki University of Technology, Publications in Computer and Information Science, Espoo, Finland, November 2000.
- [9] Dhillon I S, Modha E M, Roshan U. Diametrical clustering for identifying anti-correlated gene clusters[J]. Bioinformatics, 2003, 19(13): 1612-1619.
- [10] 张阳德. 生物信息学[M]. 北京: 科学出版社, 2005.