文章编号: 1001-0920(2009)03-0472-05

一种无监督数据驱动的学习算法

刘开第, 庞彦军, 周少玲, 栗文国
(河北工程大学 不确定性数学研究所, 河北 邯郸 056038)

摘 要: 用代表点替代类均值代表类、用加权距离替代欧氏距离作为样本与类之间的相似性度量,由此建立一种新的无监督数据聚类算法. 提取指标对分类所作贡献大小的量化值,以此为启发性知识定义加权距离,建立了用质心修正当前代表点的迭代算法. 与均值聚类等序贯算法不同,基于质心的迭代算法的批处理性可消除输入产生的随机性干扰. 采用 IRIS 数据和 Breast Cancer 数据验证了该算法的有效性.

关键词: 无监督学习;代表点;分类权;隶属度;质心驱动
中图分类号: TP391 **文献标识码**: A

Unsupervised learning algorithm based on data driving

LIU Kai-di, PANG Yan-jun, ZHOU Shao-ling, LI Wen-guo
(Institute of Uncertainty Mathematics, Hebei University of Engineering, Handan 056038, China. Correspondent: LIU Kai-di, E-mail: liukaidi@hebeu.edu.cn)

Abstract: Rep-point represents class instead of class mean, and weighted distance is regarded as the similarity measurement between sample and class instead of Euclidean distance. Then a new unsupervised data clustering algorithm is proposed. The quantized value is extracted which describes that index contributes to classification and is used as heuristic knowledge to define weighted distance. An iterative algorithm is constructed by using the center of mass to modify present rep-points. The batch mode of the iterative algorithm based on center of mass can eliminate random caused by input, which is different from sequential methods such as mean clustering algorithm. IRIS and Breast Cancer verify the effectiveness of this algorithm.

Key words: Unsupervised learning; Rep-point; Classification weight; Membership; Center of mass driving

1 引 言

随着支持向量机(SVM)[1]方法的出现,有监督模式识别的学习能力得以大幅度提高. 无监督学习则不然,因为没有供学习使用的已知分类样本,除了 N 个样本向量外没有可直接利用的分类信息,这种极弱的信息条件不利于产生新的学习算法. 10 多年来,除了 SOM(自组织聚类神经网络)[2],模糊 SOM[3], K-means 和模糊 K-means 等少数学习方法外,新的更有效的无监督学习方法十分少见. 随着计算机和检测技术的迅速发展,人们获取并存储海量数据的能力空前提高. 这种能力的提高与算法滞后的矛盾对于无监督学习尤为突出,因此对无监督学习算法的研究就显得十分重要和急迫.

SOM 算法的优点是便于用二维平面的可视性对样本进行分类;与此相伴的不足是:从高维空间到二维平面的映射不惟一,并且无法保证施行的变换是拓扑变换,因此难免造成信息失真. 这也是通常情况下 SOM 聚类效果不够理想的主要原因.

K-means 聚类是误差平方和最小意义上的最优聚类,但却存在以下不足:1) 从分类意义上讲,用均值代表类不是最优选择;2) 从分类的角度讲,样本与类之间相似性度量的选择并不合理;3) 逐个输入样本点修正类均值的序贯算法,其随机性会影响解的稳定性. 这些不足使得基于 K-means 的有关算法的学习效果和稳定性也不够理想.

鉴于上述情况,本文给出一种无监督学习的聚类算法. 用具有某种个性的代表点替代具有共性的类均值代表类;提取各维指标对分类所作贡献的量化值,以此为启发性知识定义样本点与代表点的加权距离,使之作为样本点与类之间的相似性度量;以

收稿日期: 2008-01-05; **修回日期**: 2008-04-28.
基金项目: 国家自然科学基金项目(60474019).
作者简介: 刘开第(1940—),男,山东莱州人,教授,从事复杂数据算法等研究; 庞彦军(1964—),男,河北邯郸人,教授,从事复杂数据算法等研究.

K 类隶属度为点质量的所有样本点构成的质点组的质心,修正当前的 K 类代表点,建立搜索满意类代表点的迭代算法;由此建立一种新的无监督学习算法,该算法的批处理性可消除 K-means 等序贯算法产生的随机性对解的干扰.

2　代表点聚类条件下的启发性知识获取

用 P 个代表点 $m_1(0) \sim m_P(0)$ 分别表示 P 个聚类 $C_1 \sim C_P$,这显然是一种分类. 若有两个代表点重合为一点,则两点所代表的类无法区分;如果两个代表点非常接近,则区分所代表的类是很困难的.

P 个代表点在空间中的位置由点的各维分量决定,因此空间中各维指标在分类中所起的作用(即对分类作出的贡献)大小不同. 要想对样本正确分类,就不能不考虑各指标在分类中的不同作用,即必须提取各指标对分类所作贡献的量化值作为指导样本分类的启发性知识. 为此,令

$$m_K(0) = (m_{K1}(0), m_{K2}(0), \cdots, m_{Kd}(0)), \quad K = 1 \sim P; \tag{1}$$

$$\overline{m(0)} = \frac{1}{P}\sum_{K=1}^{P} m_K(0) = (\overline{m_1}, \overline{m_2}, \cdots, \overline{m_d}); \tag{2}$$

$$\sigma_j^2(0) = \frac{1}{P}\sum_{K=1}^{P}(m_{Kj}(0) - \overline{m_j})^2, \quad j = 1 \sim d; \tag{3}$$

$$\alpha_j(0) = \sigma_j(0) \Big/ \sum_{t=1}^{d}\sigma_t(0), \quad j = 1 \sim d. \tag{4}$$

显然,式(4)定义的 $\alpha_j(0)$ 满足

$$0 \leq \alpha_j(0) \leq 1, \quad \sum_{t=1}^{d}\alpha_t(0) = 1. \tag{5}$$

称 $\alpha_j(0)$ 为以 $m_1(0) \sim m_P(0)$ 为类代表点时 j 指标的分类权.

分类权 $\alpha_j(0)$ 的物理意义是:j 指标是把 P 个类代表点区分开所作贡献大小在所有 d 个指标中所占的比例. 若 $\alpha_j(0) = 0$,则 P 个类代表点的 j 分量相同. 这时,对于区分 P 个代表点代表的类而言,j 指标已不起作用,删掉 j 指标不会影响 P 个代表点的位置,即不影响类别的划分.

3　样本点与类之间的相似性度量

用代表点代表类的目的是:如果样本点与某个类代表点更接近,则将样本点归入该代表点所代表的类. 这样就要在样本点 y_i 与代表点 $m_K(0)$ 之间定义某种距离 $d(y_i, m_K(0))$,用于表示 y_i 与 $m_K(0)$ 的接近程度,并按最小距离准则对 y_i 归类. $d(y_i, m_K(0))$ 实际上是样本点 y_i 与 C_K 类之间的相似性度量,它包含着样本 y_i 的分类信息,因此 $d(y_i, m_K(0))$ 不是单纯的欧氏距离,而是一种以 $\alpha_j(0)$ 为权的加权距离,即

$$[d(y_i, m_K(0))]^2 = \sum_{j=1}^{d}\alpha_j(0)[y_{ij} - m_{Kj}(0)]^2. \tag{6}$$

$d(y_i, m_K(0))$ 越小,样本点 y_i 属于代表点 $m_K(0)$ 代表的 C_K 类的可能性越大. y_i 属于 C_K 类的隶属度 $\mu_K(y_i)$ 定义为

$$\mu_K(y_i) \triangleq \frac{1}{\delta + d(y_i, m_K(0))} \Big/ \sum_{t=1}^{P}\frac{1}{\delta + d(y_i, m_t(0))}, \tag{7}$$

其中 $\delta > 0$ 为控制常数. 显然,由式(7)定义的隶属度满足

$$0 \leq \mu_K(y_i) \leq 1, \quad \sum_{t=1}^{P}\mu_K(y_i) = 1. \tag{8}$$

任一可能的样本点 y_i 关于任意 C_K 类都对应一个非负隶属度 $\mu_K(y_i)$. 用隶属度描述样本点与各类间的关系,比用加权距离的内容更丰富.

4　初始类代表点与质点组的质心

对于 d 维标称化空间中的 N 个样本点,将其划分为 P 个初始分类,对每个样本 $y_i (i = 1 \sim N)$ 计算[4]

$$\mathrm{sum}(y_i) = \sum_{j=1}^{d} y_{ij}, \tag{9}$$

$$M_A = \max_i \mathrm{sum}(y_i), \tag{10}$$

$$M_I = \min_i \mathrm{sum}(y_i). \tag{11}$$

如果把 N 个样本划分为 P 个类,则对每个 y_i 计算

$$1 + \frac{(P-1)[\mathrm{sum}(y_i) - M_I]}{M_A - M_I}. \tag{12}$$

若与该计算值最接近的正整数为 K,则将 y_i 归入第 K 类.

按照上述方法,可将 N 个样本划分为 P 个初始分类. 第 K 个初始分类记为 $C_K(0)$,内含 $n_K(0)$ 个样本,其均值记为 $m_K(0)$.

本文的目的是搜索 N 个样本点在空间中自然形成的样本点相对集中的 P 个区域. 由确定初始分类的作法知:作为初始代表点的 $m_K(0)$,或离第 K 个聚类区域 C_K 很近,或在第 K 个聚类区域中. 因此可认为 $m_K(0)$ 能大致代表 C_K 类,即用式(7)定义的 $\mu_K(y_i)$ 表示样本属于 C_K 类的程度具有一定的可信性.

若将 $\mu_K(y_i)$ 作为点质量赋予样本点 y_i,则能更好地代表 C_K 类的代表点,朝着以 $\mu_K(y_i)$ 为点质量的 N 个样本点构成的质点组的质心方向移动. N 个质点组的质心指出了搜索满意类代表点的搜索方向.

4.1　搜索满意类代表点的迭代算法步骤

搜索满意类代表点的迭代算法步骤如下:

1) 以初始分类的类均值作为 C_K 类的初始类代表点 $m_K(0)$, $K = 1 \sim P$. 迭代按节拍 t 进行,置 $t = 1$,设置 $t_{\max}$ 和终止常数 $\varepsilon > 0$.

2) 在当前类代表点 $m_1(t-1), m_2(t-1), \cdots, m_P(t-1)$ 的情况下,计算各指标的分类权 $\alpha_j(t-1), j=1\sim m$.

3) 计算每个样本点 y_i 到 C_K 类当前代表点 $m_K(t-1)$ 的加权距离 $d(y_i, m_K(t-1))$, $K=1\sim P, j=1\sim N$.

4) 将加权距离 $d(y_i, m_K(t-1))$ 转化为 y_i 关于 C_K 类的隶属度 $\mu_K^{(t-1)}(y_i)$, $K=1\sim P, j=1\sim N$.

5) 将隶属度 $\mu_K^{(t-1)}(y_i)$ 作为点质量赋予样本点 y_i,计算由 N 个样本点构成的质点组的质心

$$Q_K(t-1)=\left(\sum_{i=1}^{N}\mu_K^{(t-1)}(y_i)\,y_i\right)\Big/\sum_{i=1}^{N}\mu_K^{(t-1)}(y_i),\quad K=1\sim P. \tag{13}$$

6) 按下述方法修改当前类代表点:

$$m_K(t)=m_K(t-1)+\alpha(t)[Q_K(t-1)-m_K(t-1)]. \tag{14}$$

其中 $\alpha(t)$ 为步长参数,是 t 的单减函数,当 $t\to\infty$ 时,$\alpha(t)\to 0$;特殊情况下可取 $\alpha(t)\equiv 1$.

7) 计算并判断

$$J(t)=\sum_{K=1}^{P}\sum_{j=1}^{d}[m_{Kj}(t)-m_{Kj}(t-1)]^2<\varepsilon? \tag{15}$$

若答案为否,则继续;若答案为是,则转 9).

8) 判断 $t<t_{\max}$?若答案为是,则令 $t=t+1$,转 2);若答案为否,则继续.

9) 停止并输出:

①当前类代表点 $m_1(t), m_2(t), \cdots, m_P(t)$;

②当前类代表点下的分类权 $\alpha_j(t), j=1\sim d$.

4.2 算法讨论

1) 随着迭代次数的增加,相邻节拍的代表点之间的距离逐渐趋于零,因此迭代最终一定收敛.

2) 所要搜索的样本点在空间中自然形成的第 K 个相对集中的区域,实际上是无法确知的;按分量相对接近规则确定的 C_K 类,也不可能用一个代表点准确代表.因此,迭代算法中由代表点确定的 P 个聚类只能是满意聚类.

3) 与逐点输入样本的序贯算法相比,质心驱动算法对初始分类有更高的要求:每一初始分类样本的各维分量必须相对接近.

5 算法有效性检验

IRIS 数据[5] 是国际上公认的检验无监督聚类效果的典型数据. IRIS 数据分为 Ⅰ, Ⅱ, Ⅲ 3 类,每类 50 个样本,每个样本都是关于花瓣测量值的 4 维数据.

实验前,先用下述公式:

$$y_{ij}=\frac{x_{ij}-\min_i\{x_{ij}\}}{\max_i\{x_{ij}\}-\min_i\{x_{ij}\}},\quad j=1\sim 4,\ i=1\sim 150. \tag{16}$$

对 IRIS 数据进行标称化处理. 经标称化处理后,每维分量都位于闭区间[0,1]上.

实验 1 不同算法的聚类效果对比

用式(9) ~ (12) 将标称化 IRIS 数据分为 3 个初始分类,其均值 $m_1(0)$, $m_2(0)$ 和 $m_3(0)$ 分别作为初始分类的代表点. 按迭代算法搜索满意聚类点,并计算错分样本个数.

不同学习方法关于 IRIS 数据的检验结果如表 1 所示.

表 1 不同学习方法对 IRIS 数据的学习效果

学习方法	错分样本数	结果稳定性	取自文献
SOM	不少于 27	不确定	[6]
K-means	不少于 16	不确定	[6]
模糊 K-means	10 以上	不确定	[3]
Neura Gas	不少于 11	不确定	[6]
Ng-jordan	不少于 16	不确定	[6]
文献[6]	不少于 7	不确定	[6]
文献[7]	不少于 7	不确定	[7]
本 文	5	确 定	

由表 1 可知,本文算法错分样本数最少,且结果确定. 特别是该算法简捷,可重复,收敛速度快.

检验过程如下:

初始分类:Ⅰ类含样本 41 个,Ⅱ类含样本 88 个,Ⅲ类含样本 21 个. 错分样本数 44 个,其中:12 个由 Ⅰ错分到 Ⅱ,3 个由 Ⅱ错分到 Ⅰ,29 个由 Ⅲ错分到 Ⅱ.

3 个初始分类的类均值分别为

$m_1(0)=(0.1660, 0.4959, 0.0992, 0.0703)$,

$m_2(0)=(0.4684, 0.3996, 0.5487, 0.5374)$,

$m_3(0)=(0.7751, 0.4940, 0.8467, 0.8750)$.

当以均值作为初始聚类点,并按最小加权距离准则对 150 个样本重新分类时,错分样本由 44 个减少到 14 个. 这一事实说明加权距离有利于保持分类一致性.

取 $\delta=0.00005$,选择

$$\alpha(t)=\begin{cases}0.5, & t=1;\\ 1/t\sqrt{t}, & 2\le t\le t_{\max}.\end{cases}$$

则 3 次迭代后错分样本数为 5 个;一直迭代 200 次,错分样本数稳定在 5 个. 3 次迭代后聚类点分别为

$m_1(3)=(0.2211, 0.5121, 0.1599, 0.1386)$,

$m_2(3)=(0.4680, 0.3828, 0.5556, 0.5393)$,

$m_3(3)=(0.6724, 0.4534, 0.7543, 0.7749)$.

分类权为

$\alpha(3) = (0.2040, 0.0168, 0.3659, 0.4136)$.

误差为

$$\sum_{K=1}^{3}\sum_{j=1}^{4}[m_{Kj}(3) - m_{Kj}(2)]^2 = 2.3064 \times 10^{-5}.$$

当迭代 200 次后,错分样本数稳定在 5 个,误差小于 3.8×10^{-18}.

实验 2　Breast Cancer 检验数据聚类效果对比

Breast Cancer 分类数据集共 699 组数据,分为两类. 其中: 458 组属于 benign, 241 组属于 malignant. 有 16 组数据由于数据缺失而被剔除. 这样,仿真采用 683 组数据,每组数据都是 9 维数据. 其中:444 组属于 benign,239 组属于 malignant.

将 Breast Cancer 数据作为无监督学习的检验数据. 检验过程如下:

1) 将 Breast Cancer 数据按式(13) 标称化.

2) 按式(9) ～ (12) 对 683 个标称化数据进行初始分类:

第 1 类(benign) :515 个样本,其中错分样本 2 个;

第 2 类(malignant) :168 个样本,其中错分样本 73 个.

合计错分样本 2 + 73 = 75 个. 即第 1 类错分到第 2 类的 2 个,第 2 类错分到第 1 类的 73 个.

3) 两类的均值分别为

$m_1(0) =$

$(0.2736, 0.0744, 0.0919, 0.0706, 0.1476, 0.1232, 0.1564, 0.0634, 0.0151)$,

$m_2(0) =$

$(0.7163, 0.7434, 0.7189, 0.6104, 0.5569, 0.7718, 0.6250, 0.6501, 0.2262)$.

以 $m_1(0)$ 和 $m_2(0)$ 为初始代表点,计算:

①分类权

$\lambda(0) =$

$(0.0776, 0.1772, 0.1557, 0.1155, 0.0664, 0.1666, 0.0870, 0.1363, 0.0176)$.

②样本 $y_i(i = 1 \sim 683)$ 到 $m_1(0)$ 和 $m_2(0)$ 的加权距离,并按最小距离找出错分样本. 结果是:第 1 类错分到第 2 类的 6 个,第 2 类错分到第 1 类的 29 个,合计错分样本 35 个.

错分样本由 75 个减为 35 个,说明用加权距离作为样本点与类之间的相似性度量具有一定的优势.

4) 以 $m_1(0)$ 和 $m_2(0)$ 为初始类代表点,取

$$\alpha(t) = \begin{cases} 0.5, & t = 1; \\ 1/\ln t, & t \geq 2. \end{cases}$$

当迭代趋于稳定时(如 $t = 5$),代表点分别为

$m_1(5) =$

$(0.3462, 0.1919, 0.1999, 0.1653, 0.2191, 0.2231, 0.2370, 0.1658, 0.0544)$,

$m_2(5) =$

$(0.4304, 0.3012, 0.3072, 0.2537, 0.2868, 0.3616, 0.3176, 0.2631, 0.0837)$.

第 1 类错分到第 2 类的 11 个,第 2 类错分到第 1 类的 8 个,合计错分样本数 11 + 8 = 19 个. 正确分类数 664 个,平均错分率小于 2.8 %,且结果确定.

当取不同的步长参数 $\alpha(t)$ 时,达到稳定时所需的迭代次数不同. 例如:

①选择

$$\alpha(t) = \begin{cases} 0.5, & t = 1; \\ 1/\ln(t+6), & t \geq 2. \end{cases}$$

循环 70 次后,错分样本数稳定在 19 个.

②选择

$$\alpha(t) = \begin{cases} 0.5, & t = 1; \\ 1/\log_2(t), & t \geq 2. \end{cases}$$

循环 93 次后,错分样本数稳定在 19 个.

③选择

$$\alpha(t) = \begin{cases} 0.5, & t = 1; \\ 1/\sqrt{t}, & t \geq 2. \end{cases}$$

循环 125 次后,错分样本数稳定在 19 个.

Breast Cancer 检验数据的聚类效果对比如表 2 所示.

表 2　Breast Cancer 检验数据聚类效果比较

算　法	正确分类样本数	结果稳定性	取自文献
SOM	660.5 ±0.5	不确定	[6]
K-means	656.5 ±0.5	不确定	[6]
Neura Gas	656.5 ±0.5	不确定	[6]
Ng-jordan	652 ±2	不确定	[6]
文献[6]	662.5 ±0.5	不确定	[6]
本　文	664	确 定	

表 2 结果显示,本文算法错分样本数最少,且结果稳定. 该算法相对简单,可重复,具有实时性强的特点.

6　结　　论

1) 本文算法与均值聚类的根本区别在于:用具有个性的代表点替代具有共性的均值代表类. 均值聚类是用类均值代表类,以样本点与均值点的误差平方和最小作为聚类准则,每输入一个样本就要调整一次类均值,是一种序贯算法. 代表点聚类是用具有某种个性的点代表类,它是空间中 N 个样本点自然形成的 P 个聚类区域的均值. K 类质心给出了第

K 个满意代表点 $m_K(t)$ 的搜索方向,用质心修正类代表点的迭代算法是批处理算法,可消除序贯算法产生的随机性干扰.

2) 提取各指标对聚类所作贡献大小的量化值,并以此为启发性知识定义样本与代表点的加权距离,使之作为样本与类之间的相似性度量.该方法不同于其他的聚类算法,它体现了通过数据挖掘实现数据驱动的特点.

上述两条是本文算法与现有聚类算法的不同点,也是本文算法的创新点.两种典型数据的检验效果证明了本文算法的价值.

3) 无监督学习没有提供样本分类的空间信息,无论给定样本在空间如何分布,都只能把各维分量相对接近的样本归为一类作为学习规则.本文学习算法适用于在加权距离意义上各类大致呈球形分布的样本点聚类.

参考文献(References)

[1] Vladimir N Vapnik. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.

[2] Kohonen T. The self-organizing map[J]. Proc of IEEE, 1990, 78(9): 1464-1480.

[3] Bezdek J C, Tsao E C K, Pal N R. Fuzzy kohonen clustering network[C]. IEEE'92 1st Fuzzy — IEEE Proc. San Diego, 1992: 1035-1043.

[4] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2000.
(Bian Z Q, Zhang X G. Pattern recognition[M]. Beijing: Tsinghua University Press, 2000.)

[5] Everitt B S. Cluster analysi[M]. New York: Halsted Press, 1993.

[6] Francesco Camastra Alessandro Verrian. A novel kernel method for clustering[J]. IEEE Trans on Pattern and Machine Intelligence, 2005, 27(5): 801-805.

[7] Scott C. Adaptive fuzzy leader clustering complex data sets in pattern recognition[J]. IEEE Trans on Neural Networks, 1992, 3(5): 794-800.

(上接第 471 页)

[3] Zeng Lianliu, Jaroslav Svoboda. A new control scheme for nonlinear systems with disturbances[J]. IEEE Trans on Control Systems Technology, 2006, 14(1): 176-181.

[4] 房方, 谭文, 刘吉臻. 机炉协调系统的非线性输出跟踪控制[J]. 中国电机工程学报, 2005, 25(1): 147-151.
(Fang F, Tan W, Liu J Z. Nonlinear output tracking control for the coordinated system of boiler-turbing units[J]. Proc of the CSEE, 2005, 25(1): 147-151.)

[5] 孙郁松, 孙元章, 卢强, 等. 水轮机调节系统非线性 H_∞ 控制规律的研究[J]. 中国电机工程学报, 2001, 21(2): 56-60.
(Sun Y S, Sun Y Z, Lu Q, et al. Research on nonlinear robust control strategy for hydroelectric generator's valve[J]. Proc of the CSEE, 2001, 21(2): 56-60.)

[6] Euntai Kim. A fuzzy disturbance observer and its application to control[J]. IEEE Trans on Fuzzy Systems, 2002, 10(1): 77-84.

[7] Euntai Kim, Changwoo Park. Fuzzy disturbance observer approach to robust tracking control of nonlinear sampled systems with the guaranteed suboptimal H_∞ performance[J]. IEEE Trans on Systems, Man and Cybernetics, 2004, 34(3): 1574-1581.

[8] 陈谋, 姜长生, 吴庆宪. 基于干扰观测器的一类不确定非线性系统鲁棒 H_∞ 控制[J]. 控制理论与应用, 2006, 23(4): 611-614.
(Chen M, Jiang C S, Wu Q X. Robust H_∞ control of a class of nonlinear uncertain systems with disturbance observer[J]. Control Theory & Applications, 2006, 23(4): 611-614.)

[9] 刘国荣, 万百五. 一类非线性 MIMO 系统的直接自适应模糊鲁棒控制[J]. 控制理论与应用, 2002, 19(5): 693-698.
(Liu G R, Wan B W. Direct adaptive fuzzy robust control for a class of nonlinear MIMO systems[J]. Control Theory & Applications, 2002, 19(5): 693-698.)

[10] 刘国荣, 万百五. 一类非线性 MIMO 系统的间接自适应模糊鲁棒控制[J]. 控制与决策, 2002, 17(增): 676-680.
(Liu G R, Wan B W. Indirect adaptive fuzzy robust control for a class of nonlinear MIMO systems[J]. Control and Decision, 2002, 17(S): 676-680.)