

# 一种基于开方检验的特征选择方法

黄源 李茂 吕建成

(四川大学计算机科学学院 成都 610065)

**摘要** 开方检验是目前文本分类中一种常用的特征选择方法。该方法仅关注词语和类别间的关系,而没有考虑词与词之间的关联,因此选择出的特征集具有较大的冗余度。定义了词语的“剩余互信息”概念,提出了对开方检验的选择结果进行优化的方法。使用该方法可以得到既有很强表征性又有很高独立性的特征集。实验表明,该方法表现良好。

**关键词** 文本分类,特征选择,开方检验,互信息

**中图分类号** TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.5.011

## New Feature Selection Method Based on CHI

HUANG Yuan LI Mao LV Jian-cheng

(College of Computer Science, Sichuan University, Chengdu 610065, China)

**Abstract** CHI is a widely used feature selection method in text classification. This method only focuses on the relevance between features and classifications but ignores the relevance between feature and feature, resulting in a high redundancy. This paper proposed a concept about residual mutual information, and then CHI and residual mutual information were combined together to optimized the selective results. The experimental results indicate that the method is effective.

**Keywords** Text categorization, Feature selection, CHI, Mutual information

在文本分类中,特征提取对于最终分类的准确度有着巨大影响。特征提取包括两种方法:特征选择和特征抽取。特征抽取也被称为特征重参数化。特征抽取通过将原始特征空间进行变换,重新生成一个维数更小、各维之间更独立的特征空间。特征选择是从特征集  $F = \{f_1, f_2, \dots, f_n\}$  中选择一个真子集  $S = \{s_1, s_2, \dots, s_k\}$ , 满足  $k \ll n$ 。其中  $n$  为原始特征集的大小,  $k$  为选择后的特征集大小。特征选择不改变原始特征空间的性质,只是从原始特征空间中选择一部分重要的特征组成一个新的低维空间。可以看出,特征抽取涉及到语义上的分析,而目前自然语言语义处理技术尚不发达,用特征抽取方法进行特征降维的效果并不显著。相比之下,特征选择选出的特征集合是原始特征集的子集,所以更易实现,方法也更加多样,因此现阶段的文本分类多采用特征选择<sup>[1]</sup>。

在中文文本分类中,开方检验(CHI)是一种常用的并且效果很好的特征选择算法,它关注单个词与整个类别的关联度,先计算所有词语与类别的关联度,然后按照关联度从高到低选取需要的  $k$  个词语作为特征。但是该算法并没有考虑词与词之间的关联性,因此选取的特征对于整个系统虽然很有代表性但会出现很多冗余的词语。这些冗余的词语不仅增加了选择的特征集的维度,也降低了分类精度。为了剔除冗余特征,在选取过程中引入了互信息,提出了“剩余互信息”的概念,定义了词的“剩余互信息”用于选择冗余度低的词。本文在特征选取过程中先用开方检验进行第一次特征选择,然后

再根据词的“剩余互信息”对特征集进行优化,剔除冗余特征,降低特征维度,提高分类精度。从实验结果来看,该方法取得了不错的效果。

## 1 相关工作

### 1.1 特征的关联性和冗余性

通常情况下,特征选择的研究都考虑类别和词语间的关联性,很少关注特征的冗余性。根据特征和类别间的关联性的强弱,John, Kohavi 和 Pfleger 将特征分为 3 种类型:强相关、弱相关和不相关<sup>[2]</sup>。假设集合  $F$  为包含所有特征的特征集,  $f_i \in F$  是一个特征,集合  $S_i = F - \{f_i\}$ ,  $C$  为类别,那么上述 3 类特征可以用下面的方式定义。

**定义 1(强相关)** 如果一个特征  $f_i$  是一个强相关特征,那么它的充要条件为

$$P(C|f_i, S_i) \neq P(C|S_i)$$

**定义 2(弱相关)** 如果一个特征  $f_i$  是一个弱相关特征,那么它的充要条件为

$$P(C|f_i, S_i) = P(C|S_i)$$

并且  $\exists S_i' \subset S_i$ , 使得  $P(C|f_i, S_i') \neq P(C|S_i')$ 。

**推论(不相关)** 如果一个特征  $f_i$  是一个不相关特征,那么它的充要条件为

$$\forall S_i' \subset S_i, P(C|f_i, S_i') = P(C|S_i')$$

如果一个特征具有强相关性,那么表明这个特征是必须

到稿日期:2014-02-12 返修日期:2014-04-15 本文受教育部博士点基金资助项目(2010081110053)资助。

黄源(1989—),男,硕士生,主要研究方向为数据挖掘, E-mail: huangyuan1219@gmail.com; 李茂(1974—),男,博士生,主要研究方向为数字艺术, E-mail: 457249960@qq.com; 吕建成 男,博士,教授,博士生导师,主要研究方向为机器学习、深度学习、神经网络学习算法及其应用, E-mail: lvjiancheng@gmail.com。

的,在进行分类判定时它是不可或缺的。如果一个特征具有弱相关性,那么表明这个特征并非总是必要的,但是在一定条件下最优特征集中必须包含它。如果一个特征属于不相关特征,那么说明这个特征在任何时候都是没有用处的。一个最优特征集应当包含所有强相关特征和部分弱相关特征。但是上面的定义并没有说明应当选择哪些弱相关特征。因此,下面再谈谈特征的冗余性。

假设:在一个类别  $C$  中的所有文章中都出现了词  $A$ ,同时其他类别中的文章都没有出现词  $A$ ,另外在类别  $C$  的 90% 的文章中出现了词  $B$ ,而且其他类别中的文章也都没有出现词  $B$ ,那么单独看词  $A$  和词  $B$  都具有很好的表征性。但仔细观察发现,词  $B$  总是随着词  $A$  出现的,即包含  $A$  的文章肯定包含  $B$ ,但包含  $B$  的文章不一定包含  $A$ ,那么词  $B$  其实是冗余的,最优特征集应当包含  $A$ ,但不应包含  $B$ 。那么如何来判定一个词的冗余度呢,可以考虑使用互信息 (Mutual Information) 来衡量。互信息是信息论里一种有用的信息度量,指两个事件集合之间的相关性。设有两个随机变量  $X$  和  $Y$ ,它们之间的联合概率为  $p(x, y)$ ,边缘概率分别为  $p(x)$  和  $p(y)$ ,那么  $X$  和  $Y$  之间的互信息值为:

$$I(X; Y) = H(X) - H(X|Y) \quad (1)$$

其中,  $H()$  函数表示信息熵。通过上面的式子可以计算出两个特征间的互信息值,两个词关系越紧密,它们的互信息值也就越大。比如词  $A$  是强相关特征,词  $B, C$  是弱相关特征,  $A$  和  $B$  的互信息值很高,  $A$  和  $C$  的互信息值很低,那么,在  $A$  已经被选取的情况下,我们认为  $B$  比  $C$  的冗余度更高,因此在特征选择时倾向于选取  $C$  而不是  $B$ 。具体的方法会在第 2 节进行详细描述。

## 1.2 开方检验 (CHI)

CHI 是一种用途很广的用于检验两个变量独立性的方法。它的基本的思想就是通过观察实际值与理论值的偏差来确定理论的正确与否。在一般应用中,我们先做一个虚无假设:一个样本中已发生事件的次数服从某个特定的理论分配。通常假设两个变量是独立的,然后判断观察值与理论值之间差异程度的大小,如果两者相近,我们就认为偏差是由于自然误差导致的,这两个变量是相互独立的,即认为虚无假设正确;如果两者之间差异过大,我们就认为这两个变量是相互关联的,即拒绝虚无假设,而接受备择假设<sup>[3]</sup>。

假设理论值为  $E$ , 实际值为  $x$ ,  $x$  的观察值有  $n$  个即  $x_1, x_2, \dots, x_n$ , 那么衡量差异度的公式为:

$$\sum_{i=1}^n \frac{(x_i - E)^2}{E} \quad (2)$$

通过式(2)可以算出一个值,再用这个值与设定好的阈值做对比,如果该值比阈值大,就认为虚无假设不成立;如果该值比阈值小,则认为虚无假设成立。

将 CHI 运用于文本分类的特征选择时稍有不同,由于很难说一个词与类别之间的开方值达到多少时这个词才能表征这个类别,只通过 CHI 方法来选出一些最相关的词即可。一个特征  $f_i$  的 CHI 权重计算公式如下:

$$x^2(f_i, C_j) = \frac{n[P(f_i, C_j)P(\bar{f}_i, \bar{C}_j) - P(f_i, \bar{C}_j)P(\bar{f}_i, C_j)]^2}{P(f_i)P(C_j)P(\bar{f}_i)P(\bar{C}_j)} \quad (3)$$

其中,  $n$  为总的文本数,  $P(f_i)$  表示一个训练文本中包含特征  $f_i$  的概率,  $P(C_j)$  表示一个训练文本的类别为  $C_j$  的概率,

$P(f_i, C_j)$  表示一个训练文本包含特征  $f_i$  且它的类别为  $C_j$  的概率,  $P(\bar{f}_i, \bar{C}_j)$  表示一个训练文本不包含特征  $f_i$  同时它的类别不为  $C_j$  的概率,  $P(f_i, \bar{C}_j)$  表示样本集中包含特征  $f_i$  且样本不属于类别  $C_j$  的概率,  $P(\bar{f}_i, C_j)$  表示样本集中不包含特征  $f_i$  且样本属于类别  $C_j$  的概率。CHI 值越大,说明特征  $f_i$  和类别  $C_j$  就越相关,最后按照 CHI 值从大到小排序,按需求选取一定数量的词作为最终选取的特征放入特征集中。大量实验和应用表明,CHI 在中文文本分类中有着不错的表现。

## 2 新的特征提取方法

CHI 虽然有着不俗的表现,但由于其只关注特征与类别间的联系,没有考虑特征之间的联系,因此选取出来的特征冗余度高。剔除掉这些特征不仅可以降低特征集的维度,减少分类时的计算量,并且通过实验发现,分类精度还会有一定程度的提升。

之前提到用互信息来衡量一个词的冗余度,那么在具体选取过程中应该如何使用互信息呢? 这里定义一个“剩余互信息”的概念,它表示在除去已被选择的特征的情况下一个词和系统间的关联度的高低,用公式表示如下:

$$I(C; f_i) - \sum_{f_s \in S} I(f_s; f_i) \quad (4)$$

其中,  $f_i$  表示尚未被选取的特征,集合  $S$  为已被选取的特征的集合,  $f_s$  即为已被选择的特征。一个词的剩余互信息包含两部分:这个词与整个系统间的互信息  $I(C; f_i)$  和这个词与已经被选择的特征间的互信息  $\sum_{f_s \in S} I(f_s; f_i)$ 。在选择过程中,我们优先选择剩余互信息值高的词。从式(4)可以看出,若一个候选特征  $f_i$  与整个系统有较高的互信息值,但如果它与已选取的特征  $f_s \in S$  有很高的关联性 ( $\sum_{f_s \in S} I(f_s; f_i)$  很高),也会导致该特征的剩余互信息值偏低,那么这个候选特征  $f_i$  并不会被优先选取。这样选择出来的特征不仅有很好的表征性,特征间也具有较低的冗余度。

我们提出的新的特征选择算法的具体流程如下。

假设有  $m$  个类别  $C_j, j=1, 2, \dots, m$ , 共有  $n$  个初始特征  $f_i \in F, i=1, 2, \dots, n$ 。首先,采用 CHI 方法选取出所需的  $k$  个特征  $f_i' \in F'$ , 然后分别计算所有特征  $f_i' \in F'$  和整个系统的互信息值  $I(C; f_i')$ :

$$I(C; f_i') = H(C) - H(C|f_i') = -\sum_{k=1}^m P(C_j) \log_2 P(C_j) + P(f_i') \sum_{k=1}^m P(C_j | f_i') \log_2 P(C_j | f_i') + P(\bar{f}_i') \sum_{k=1}^m P(C_j | \bar{f}_i') \log_2 P(C_j | \bar{f}_i') \quad (5)$$

其中,  $P(C_j | f_i')$  表示如果特征  $f_i'$  存在于一个训练样本中,那么这个样本的类别为  $C_j$  的概率;  $P(C_j | \bar{f}_i')$  表示训练样本中如果一个文本不包含特征  $f_i'$ , 它的类别为  $C_j$  的概率。从这  $k$  个特征中选择使  $I(C; f_i')$  取最大值的特征作为第一个被选择特征,放入特征集  $S$  (初始为空集)中,并把把这个特征从集合  $F$  中除去,然后计算集合  $F$  中剩下的  $k-1$  个特征的剩余互信息值:

$$I(C; f_i') - \sum_{f_s \in S} I(f_s; f_i') \quad (6)$$

再从这  $k-1$  个特征中选择剩余互信息值最大的特征放入集合  $S$  中,并将这个特征从集合  $F$  中除去,接下来重复上述过程,直到选出所需的  $k'$  ( $k' < k$ ) 个特征,这  $k'$  个特征即为最终的选取结果。这样每次选取的特征不仅有着很好的表征

性,而且和已选取的特征之间有着较高独立性,然后便得到一个比原来维度更低但精度更高的特征集。

算法具体描述如下:

给定一个包含  $n$  个特征的原始特征集合  $F$ , 找出一个包含  $k'$  个特征的特征子集  $S \subset F$ , 集合  $S$  中的特征既有很强的表征性, 又有较低的冗余度。具体算法描述如下:

- 1) 初始化:  $F \leftarrow n$  个初始特征,  $S \leftarrow$  空集,  $F' \leftarrow$  空集;
- 2) 使用 CHI 方法进行第一次特征选择, 选出权值最大的  $k$  个特征放入集合  $F'$  中;
- 3) 计算每个特征  $f_i' \in F'$  与整个系统的互信息值  $I(C; f_i')$ ;
- 4) 找出具有最大  $I(C; f_i')$  值的特征  $f_i'$ , 然后更新集合  $F'$  和  $S: F' \leftarrow F' \setminus \{f_i'\}, S \leftarrow \{f_i'\}$ ;
- 5) 采用贪心算法进行选取, 重复以下步骤直到  $|S| = k'$ :
  - a) 计算所有候选特征  $f_i' \in F'$  的剩余互信息值;
  - b) 选择剩余互信息值最大的特征  $f_i' \in F'$ , 然后更新集合  $F'$  和  $S: F' \leftarrow F' \setminus \{f_i'\}, S \leftarrow \{f_i'\}$ ;
- 6) 输出结果: 包含  $k'$  个被选取的特征的集合  $S$ 。

### 3 实验

本文实验使用了搜狗实验室的文本分类语料库。分类算法使用的是 SVM 算法。

实验 1 首先选取 4 个类进行实验, 每个类选取 500 个文本作为训练样本, 然后每个类另外选取 500 个文本作测试。先使用 CHI 选取 1000 个特征, 然后使用新方法在此基础上选取了 900 个特征。另外再使用 CHI 直接选取 900 个特征。比较这 3 种情况下分类的准确度, 结果如表 1 所列。

表 1 实验 1 结果

特征选择方法	准确度
CHI 选择 1000 个特征	89.4%
CHI 选择 900 个特征	90.2%
新方法选择 900 个特征	91.94%

实验 2 类别依然是 4 个, 每个类选取 500 个文本作为训练样本, 然后每个类另外选取 500 个文本作测试。使用 CHI 选取 1500 个特征, 然后使用新方法在此基础上选取了 700 个特征。另外再使用 CHI 直接选取 700 个特征。比较这 3 种情况下分类的准确度, 结果如表 2 所列。

表 2 实验 2 结果

特征选择方法	准确度
CHI 选择 1500 个特征	87.45%
CHI 选择 700 个特征	90.14%
新方法选择 700 个特征	90.4%

从上面两个实验结果可以看出, 使用新方法略优于 CHI 算法, 效果不明显主要是由于类别不多, 冗余特征带来的干扰还不大。另外第 2 个实验中虽然用 CHI 直接选取 700 个特征和新方法选择 700 个特征的特征最终分类准确度相当, 但其实它们之间仅有 149 个特征相同, 而后的特征间的独立性更高。

实验 3 将类别增加到 6 个, 每个类选取 500 个文本作为训练样本, 然后每个类另外选取 500 个文本作测试。先使用 CHI 选取 1000 个特征, 然后使用新方法在此基础上选取了 800 个特征, 另外再使用 CHI 直接选取 800 个特征。比较这 3 种情况下分类的准确度, 结果如表 3 所列。

表 3 实验 3 结果

特征选择方法	准确度
CHI 选择 1000 个特征	73.05%
CHI 选择 800 个特征	75.4%
新方法选择 800 个特征	85.0%

实验 4 最后使用了 8 个类, 每个类选取 500 个文本作为训练样本, 然后每个类另外选取 500 个文本作测试。先使用 CHI 选取 1000 个特征, 然后使用新方法在此基础上选取了 900 个特征, 另外再使用 CHI 直接选取 900 个特征。比较这 3 种情况下分类的准确度, 结果如表 4 所列。

表 4 实验 4 结果

特征选择方法	准确度
CHI 选择 1000 个特征	74.5%
CHI 选择 900 个特征	62.7%
新方法选择 900 个特征	80.16%

从实验 3 和实验 4 可以看到, 在类别增加后, 直接采用 CHI 方法选择的特征进行分类的准确度大幅下降, 而采用新方法依然保持着较高的准确度。

实验 4 中用 CHI 方法选取的 900 个特征和用新方法选取的 900 个特征权重排名前 15 的特征的对比结果如表 5 所列。

表 5 新方法和 CHI 选择的特征对比

CHI	新方法
www.sogou.com	www.sogou.com
搜狗	比赛
网页	利害关系
搜索	作战
找到	考试
比赛	旅游
本人所	用户
文责自负	图库
利害关系	治疗
入市	学生
个人观点	公司
关系人	报道
利害	景点
所述	来源
仅供参考	药品

从表 5 可以看出, CHI 选取的词之间关联性很强, 一些词显得冗余, 而新方法优先选取的词之间有较强独立性。

从上面的实验可以看出, 使用新方法不仅能降低特征维度, 而且能提高分类准确度, 特别在类别较多时效果更明显。但目前面临的问题是, 在对 CHI 选取的特征集优化时, 选择多少个特征最合适。

**结束语** 本文通过引入词的剩余互信息对 CHI 的特征选择结果进行优化, 降低了选择的特征集的冗余度, 减少了分类计算的复杂度, 提高了分类的准确度, 并且该方法简单易行, 适用于所有文本分类。

但是, 该方法还有很多尚未解决的问题, 比如在进行优化时, 选择多少个特征合适, 这将是下一步的主要研究方向。

### 参考文献

- [1] 胡洁. 高维数据特征降维研究综述[J]. 计算机应用研究, 2008, 25(9): 2601-2606
- [2] John H, Kohavi R, Pflieger K. Irrelevant feature and the subset selection problem[C]// Proc. of the 11th Int. Conf. on Machine Learning, 1994. San Francisco: Morgan Kaufmann Publishers, 1994: 121-129

(下转第 77 页)

校正系数,与其他的滤波算法相比,其能够很好地保持图像的细节信息,给人以良好的主观感受。在噪声方差、噪声均值以及信噪比这3个主要的性能指标上,本算法有明显的改进。

**结束语** 本文提出的多层自适应形态滤波算法利用多层结构变换形式的优点,在层与层之间灵活地选择不同结构元素计算得到的中间结果,以及优化设置一维偏倚校正系数向量来降低偏倚现象对滤波产生的负面影响。这样使得该算法在滤除不同噪声和保护图像细节方面相对现有的滤波算法均有较大的改善,提高了形态滤波算法的适应性。在仿真过程中发现,初始结构元素设置越多,对于图像的描述性就越强,滤波后输出的结果就越好。但是,这样会导致程序的时间复杂度和空间复杂度会相应提高。因此,下一步的工作将结合实际图像的特点,在设置初始结构元素时进一步考虑对结构元素进行选择的方法,达到在保持滤波效果不变的同时降低时间复杂度和空间复杂度的目的。

### 参考文献

- [1] 徐艳蕾,赵继印,李敏.基于顺序形态学的医学图像插值算法的研究[J].电子学报,2010,38(5):32-37
- [2] Adesso P, Conte R, Longo M. A pansharpening algorithm based on genetic optimization of Morphological Filters,2012[C]//IEEE; Piscataway, NJ,2012;5438-5441
- [3] Song H, Huang B. Shadow Detection and Reconstruction in High Resolution Satellite Images via Morphological Filtering and Example-Based Learning[J]. IEEE Transactions on Geoscience and Remote Sensing,2013,23(99):1-10
- [4] Ashtiani S,Shahrtash M. Feature-oriented De-noising of Partial Discharge Signals Employing Mathematical Morphology Filters [J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2012,19(6):2128-2136
- [5] Bouaynaya N, Charif-Chefchaoui M. M-Idempotent and Self-Dual Morphological Filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(4):805-813
- [6] Bouaynaya N, Charif-Chefchaoui M. M-Idempotent and Self-Dual Morphological Filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2012,34(4):805-813
- [7] Guan Ming-shan, Ren Hong, Ma Yan. Multi-scale morphological filtering method for preserving the details of images[C]//IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA). IEEE; Piscataway, NJ, 2009; 417-420
- [8] 白银刚,于盛林.一种广义多结构元形态滤波器[J].电子学报,2009,37(4):792-796
- [9] 余农,吴常泳,汤心溢.红外目标检测的自适应背景感知算法[J].电子学报,2005,33(2):200-204
- [10] 张伟,王军锋,王涛.一种基于改进算子的形态学边缘检测算法[J].计算机技术与发展,2013,28(4):24-28
- [11] Verd-Monedero R,Angulo J,Serra J. Anisotropic Morphological Filters With Spatially Variant Structuring Elements Based on Image Dependent Gradient Fields [J]. IEEE Transactions on Image Processing,2011,20(1):200-212
- [12] Guan Jing, Zhang Tian-xu, Wang Xiao-ping. New class of Grayscale Morphological Filter to enhance infrared building target[J]. Aerospace and Electronic Systems Magazine,2012,27(6):5-10
- [13] Tan Xiang-yu, Yang Zhuo, Zhao Xian-ping. Applied research of multilevel generalized morphological filter for ultrasonic signal of PD in processing,2011[C]//IEEE; Piscataway, NJ,2011:297-300
- [14] Li Qi-Yue,Xu Jie,Wang Wei-hua. Slope displacement prediction based on morphological filtering[J]. Journal of Central South University,2013,20(6):1724-1730
- [15] Lei Tao, Fan Yang-yu. Double operator morphological filters [J]. Zidonghua Xuebao/Acta Automatica Sinica,2011,37(4):449-463
- [16] 朱康,贺新光.基于形态学和 Contourlet 系数区域特征的遥感图像融合方法[J].计算机科学,2013,40(4):301-305
- [17] Li jin,Tang Jing-tian,Xiao Xiao. De-noising algorithm for magnetotelluric signal based of mathematical morphology filtering [J]. Noise and Vibration Worldwide,2011,42(11):65-72
- [18] Hung Mao-hsiung, Pan Jeng-shyang, Hsieh Chaur-heh. A fast algorithm of temporal media filter for background subtraction [J]. Journal of information Hiding and Multimedia Signal Processing,2014,5(1):33-40
- [19] Jasper. 文本分类入门(十)特征选择算法之开方检验[OL]. 2008-08-31[2014-01-18]. <http://www.blogjava.net/zhenandaci/archive/2008/08/31/225966.html>
- [4] Yu Lei, Liu Huan. Efficient Feature Selection via Analysis of Relevance and Redundancy[J]. Journal of Machine Learning Research,2004,10(5):1205-1224
- [5] Battiti R. Using mutual information for selecting features in supervised neural net learning[J]. IEEE Trans. Neural Network, 1994,5(4):537-550
- [6] Estevez P, Tesmer M, Perez C, et al. Normalized mutual information feature selection [J]. IEEE Trans. Neural Network, 2009,20(2):189-201
- [7] Sun X, Liu Y, Xu M, et al. Feature selection using dynamic weights for classification[J]. Knowledge-Based Systems,2013,37:541-549
- [8] Li B, Chow T W S, Huang D. A novel feature selection method and its application[J]. Journal of Intelligent Information Systems,2013,41(2):235-268
- [9] Lee S, Park Y T, d'Auriol B J. A novel feature selection method based on normalized mutual information [J]. Applied Intelligence,2012,37(1):100-120
- [10] Aliferis C F, Statnikov A, Tsamardinos I, et al. Local causal and markov blanket induction for causal discovery and feature selection for classification part i: Algorithms and empirical evaluation [J]. The Journal of Machine Learning Research,2010,11:171-234
- [11] Lv Jian-cheng, Tan K K, Zhang Yi, et al. A family of fuzzy learning algorithms for robust principal component analysis neural networks[J]. IEEE Transactions on Fuzzy Systems,2010,18(1):217-226
- [12] Lv Jian-cheng, Zhang Yi, Tan K K. Global Convergence of Oja's PCA Learning Algorithm with a Non-Zero-Approaching Adaptive Learning Rate[J]. Theoretical Computer Science,2006,367(3):286-307

(上接第56页)