第 25 卷 第 9 期
Vol. 25 No. 9

控 制 与 决 策
Control and Decision

2010 年 9 月
Sep. 2010

文章编号:1001-0920(2010)09-1435-06

一种多元核Logistic回归说话人辨别方法

郑建炜[a], 王万良[a], 王震宇[b], 蒋一波[a]

(浙江工业大学 a. 计算机科学与技术学院，b. 信息与工程学院，杭州 310023)

摘 要: 针对文本无关话者辨别多分类目标和大训练样本情况, 将经典Logistic回归模型进行多元化变形, 并叠加 L_2 惩罚因子以提高模型泛化能力. 将最优目标负对数Logistic公式对偶化, 并利用序列最小优化算法进行模型训练, 速率优于传统多元核Logistic回归训练算法. 实验显示, 该模型构建简单, 训练算法快捷, 且识别率优于经典支持向量机与二元核Logistic回归模型所生成的“一对一”多分类方法.

关键词: Logistic 回归；序列最小优化；话者辨别；核技巧

中图分类号: TP391.4 **文献标识码:** A

Speaker identification based on multi-class kernel Logistic regression model

ZHENG Jiang-wei[a], *WANG Wan-liang*[a], *WANG Zhen-yu*[b], *JIANG Yi-bo*[a]

(a. College of Computer Science and Technology，b. College of Information Engineering，Zhejiang University of Technology，Hangzhou 310023，China．Correspondent: WANG Wan-liang，E-mail：wwl@zjut.edu.cn)

Abstract: The traditional Logistic regression model is transformed to multi-class kernel Logistic model applying for text-independent speaker identification, which is nonlinear and more than just two classes. The penalty factor is added for enhancing model generalization ability. Then an iterative algorithm is proposed based on the solution of a dual problem by using ideas similar to those of the sequential minimal optimization algorithm for support vector machines. Experiments show that the algorithm is robust and fast, and the recognition rate is as good as widely used methods such as SVM while being used in text-independent speaker identification.

Key words: Logistic regression；Sequential minimal optimization；Speaker recognition；Kernel trick

1 引 言

与文本无关的话者辨别是指通过对说话人语音信号的分析处理, 自动确认说话人是否在所记录的话者集合中, 继之确认说话人的具体身份. 随着计算机技术、信号处理技术的发展, 话者辨识研究已受到极大的关注, 获得了广泛的应用. 如说话人跟踪系统、网络登入身份验证等场合.

识别模型构建是话者辨别系统性能优劣的关键技术之一. 传统的话者辨识模型包括混合高斯、隐马尔可夫等产生性模型[1]. 这些模型虽然能获得极高的识别效率, 但需要大量的训练样本来优化模型参数. 随着基于核技巧的分辨性模型, 如支持向量机的产生并应用[2], 说话人辨别在获得更高的识别率的同时还降低了对训练样本量的要求.

核Logistic回归是另一种高效辨别性分类器[3], 主要用于生成分类判别中的后验概率[4], 已经成功应用于基因病理选择[5]、信用卡风险度分类[6]、孤立字识别[7]等场合. 对于说话人辨别技术而言, 尽管支持向量机已体现了强大的竞争力, 但它固有的缺陷也很明显, 如二元决策非概率输出、多分类扩展能力差等. 而核Logistic回归具有天然的后验概率输出, 以及对多元分类良好的扩展性, 这使得它在说话人辨别这些多分类判别情形中占据一定的优势. 已有的核Logistic回归应用于说话人辨别技术[8], 仅仅对二元模型进行简单应用, 虽然识别率优于经典算法, 但模型构建复杂, 且没有体现Logistic回归多元化扩展应用的优势. 本文直接将多元核Logistic回归应用于说话人辨别, 并针对语音样本数量众多的情况, 提出

收稿日期: 2009-08-23；**修回日期:** 2009-10-14.

基金项目: 国家自然科学基金项目(60573123).

作者简介: 郑建炜(1983−), 男, 杭州人, 博士, 从事模式识别、计算机视觉等研究；王万良(1957−), 男, 杭州人, 教授, 博士生导师, 从事计算机控制与智能自动化、人工智能等研究.

快速训练算法,所构建的模型响应速度快、识别率高,且更加直观.

2 核 Logistic 回归

说话人辨别应用中,设定需识别的说话人数为 K,给定训练样本集为 $\{\bar{x}_1, c_1\}, \{\bar{x}_2, c_2\}, \cdots, \{\bar{x}_n, c_n\}$. 其中: 输入 $\bar{x}_i$ 为 p 维说话人特征向量, 即 $\bar{x}_i = (x_{i1}, x_{i2}, \cdots, x_{ip})^{\mathrm{T}}$; 输出样本标签 c_i 为有限集 $\{1, 2, \cdots, K\}$ 之一. 则训练目标为优化判别模型, 使得当给定新的一段语音信号的若干帧 $\bar{x}$ 参数时, 分类器能从 $\{1, 2, \cdots, K\}$ 中选取一个正确的类别标签, 即确定相应说话人身份.

经典的 Logistic 回归模型为二元分类器, 即 $K = 2$, 定义线性判别函数[9]为

$$g(\bar{x}_i) = \mathrm{logit}(p(\bar{x}), \beta) = \log \frac{p(c_i = 1)}{p(c_i = 0)} = \beta^{\mathrm{T}} \bar{x}_i + \beta_0, \tag{1}$$

则可得样本隶属于类别 1 的后验概率

$$p(c_i = 1|\bar{x}) = \frac{\exp(g(\bar{x}))}{1 + \exp(g(\bar{x}))}. \tag{2}$$

Logistic 回归问题是将线性函数 $g(\bar{x}_i) = \beta^{\mathrm{T}} \bar{x}_i + \beta_0$ 的参数 β 最优化的问题. 假设样本目标标签 $c_i \in \{0, 1\}$ 依据输入样本集 X 服从 Bernoulli 分布, 样本似然度

$$l(\beta) = \prod_{i=1}^{n} (p(c_i = 1|\bar{x}_i)^{c_i} (1 - p(c_i = 1|\bar{x}_i)))^{1 - c_i}. \tag{3}$$

对其取负对数, 得到线性 Logistic 回归模型的原始最小化目标泛函为

$$L(\beta) = -\sum_{i=1}^{n} \left[c_i g(\bar{x}_i) - \log(1 + \exp g(\bar{x}_i)) \right]. \tag{4}$$

针对说话人辨别输入特征参数的线性不可分性, 必须将线性回归模型扩展为非线性回归模型, 从而得到核 Logistic 回归. 通过非线性映射 $\Phi : R^p \to F$ 将原输入空间映射到高维特征空间 F. 在 F 空间中, β 可表示为[10]

$$\beta = \sum_{i=1}^{n} \alpha_i \Phi(\bar{x}_i). \tag{5}$$

高维空间中构建 $g'(\bar{x}_i) = \beta^{\mathrm{T}} \Phi(\bar{x}_i) + \beta_0$, 即

$$g'(\bar{x}) = \beta^{\mathrm{T}} \Phi(\bar{x}) + \beta_0 = \Big(\sum_{i=1}^{n} \alpha_i \Phi(\bar{x}_i) \Big) \Phi(\bar{x}) + \beta_0 = \sum_{i=1}^{n} \alpha_i K(\bar{x}_i, \bar{x}) + \beta_0. \tag{6}$$

这样, 对于原输入空间而言, $g'(\bar{x})$ 就成为一个非线性函数, 其中 $K(x, y)$ 为满足 Mercer 条件的核函数. 应用较为广泛的核函数有径向基核函数

$$K(x, y) = \exp\Big(- \frac{||x - y||^2}{\sigma} \Big)$$

和 m 阶多项式核函数

$$K(x, y) = (x^{\mathrm{T}} \times \sigma \times y + 1)^m.$$

利用核技巧, 核化的后验概率为

$$p(c_i = 1|\bar{x}) = \frac{1}{1 + \exp\Big(- \sum_{i=1}^{n} \alpha_i K(\bar{x}_i, \bar{x}) - \beta_0 \Big)}.$$

而核 Logistic 回归的目标泛函为

$$\min L(\alpha) = -\sum_{i=1}^{n} \sum_{j=1}^{n} [c_i \alpha_j K(\bar{x}_i, \bar{x}_j) + \beta_0] + \sum_{i=1}^{n} \log(1 + \exp[\alpha_j K(\bar{x}_i, \bar{x}_j) + \beta_0]). \tag{7}$$

3 多元核 Logistic 回归

说话人辨别系统中一般 $K > 2$, 如果直接应用二元核 Logistic 模型进行辩识, 则需要采用“一对多”或“一对一”方法构建多个分类器, 这样增加了模型构建的繁琐度. 事实上, 核 Logistic 回归可很自然地进行多元分类扩展, 即

$$p(c_i = k|\bar{x}; \beta) = \frac{\exp({\beta_k}^{\mathrm{T}} \bar{x} + \beta_{k0})}{\sum_{j=1}^{K} \exp({\beta_j}^{\mathrm{T}} \bar{x} + \beta_{j0})}. \tag{8}$$

其中: $k = 1, 2, \cdots, K$; $\beta = [{\beta_1}^{\mathrm{T}}, {\beta_2}^{\mathrm{T}}, \cdots, {\beta_K}^{\mathrm{T}}]$, $\beta \in R^{p \times K}$ 为最优化参数. 对于新输入向量 $\bar{x}$, 分类结果为

$$\arg\max_{k \in \{1, 2, \cdots, K\}} (p(c_i = k|\bar{x}; \beta)). \tag{9}$$

最优参数的推导依然通过最小化负对数似然度

$$\min_{\beta} l(\beta) = -\log \Big(\prod_{i=1}^{n} p(c_i = k|\bar{x}_i; \beta) \Big) = \sum_{k=1}^{K} \sum_{c_i = k} \Big[-{\beta_k}^{\mathrm{T}} \bar{x}_i - \beta_{k0} + \log \Big(\sum_{j=1}^{K} \exp({\beta_j}^{\mathrm{T}} \bar{x} + \beta_{j0}) \Big) \Big]. \tag{10}$$

为增强模型的泛化能力, 对最优泛函叠加 L_2 正则化部分. 最终多元 Logistic 回归模型目标泛函为

$$\min H = \sum_{i=1}^{n} \Big[-\bar{c}_i^{\mathrm{T}} {\beta_k}^{\mathrm{T}} x_i - \beta_{k0} + \log \Big(\sum_{j=1}^{K} \exp({\beta_j}^{\mathrm{T}} \bar{x} + \beta_{j0}) \Big) \Big] + \frac{\lambda}{2} \sum_{k=1}^{K} ||\beta_k||^2. \tag{11}$$

其中: $\bar{c}_i$ 为 K 维向量, 当 $c_i = k$ 时, $\bar{c}_i = (0, 0, \cdots, 1, \cdots, 0)$, 1 为第 k 维. 同样, 应用核技巧, 相应的多元核 Logistic 回归模型目标泛函为

$$\min H' = \sum_{i=1}^{n} \sum_{j=1}^{n} [-\bar{c}_i^{\mathrm{T}} \alpha_{jk} K(\bar{x}_i, \bar{x}_j) - \beta_{k0}] + \sum_{i=1}^{n} \log \Big(\sum_{j=1}^{K} \exp \Big(\sum_{m=1}^{n} \alpha_{mj} K(\bar{x}_m, \bar{x}_i) +$$

$$\beta_{j0})\Big)+\frac{\lambda}{2}\sum_{k=1}^{K}\sum_{i,i'}\alpha_{ik}\alpha_{i'k}K(\bar{x}_i,\bar{x}_{i'}). \tag{12}$$

4 快速对偶训练

核 Logistic 回归模型的训练算法很多, 有迭代重加权最小平方法 (IRRLS)[11], 牛顿-拉斐逊法 (NL), 共轭梯度法 (CG), 信任区域牛顿法 (TRNM) 等. 所有这些方法都在一定程度上提升了训练速度, 但在训练样本数量众多且多分类目标的场合(如说话人辨别中), 每次迭代过程仍带有相当计算量的矩阵逆操作. Keerthi 等[12]首次将支持向量机快速训练算法最小序列优化应用于二元核 Logistic 回归模型训练中, 在每次迭代过程中只优化两个 α 系数, 避免了费时的矩阵操作. 为了应用于说话人辨别多元分类场合, 本文将最小序列优化扩展为多元分类训练算法.

4.1 目标泛函对偶化

原目标泛函 (11) 等同于

$$\min H'=C\sum_{i=1}^{n}\sum_{k=1}^{K}g(\xi_{ik})+\frac{1}{2}\sum_{k=1}^{K}||\bar{\beta}||^2. \tag{13}$$

其中: $C=1/\lambda$ 为正则化常量, 且

$$\xi_{ik}=\beta_{k0}+\bar{\beta}_k^{\rm T}\bar{x}_i, \tag{14}$$

$$g(\xi_{ik})=-c_{ik}\xi_{ik}+\log({\rm e}^{\xi_{i1}}+\cdots+{\rm e}^{\xi_{iK}}). \tag{15}$$

将其转化为拉格朗日形式为

$$\min\widehat{H}=$$
$$C\sum_{i=1}^{n}\sum_{k=1}^{K}g(\xi_{ik})+\frac{1}{2}\sum_{k=1}^{K}||\bar{\beta}_k||^2+\alpha_0\sum_{k=1}^{K}\beta_{k0}+$$
$$\sum_{i=1}^{n}\sum_{k=1}^{K}\alpha_{ik}(\xi_{ik}-\beta_{k0}-\bar{\beta}_k^{\rm T}\bar{x}_i). \tag{16}$$

其中: α_{ik},α_0 为拉格朗日乘子, 则 KKT 条件如下:

$$\partial\widehat{H}/\partial\beta_{k0}=\alpha_0-\sum_{i=1}^{n}\alpha_{ik}=0,\forall k; \tag{17}$$

$$\partial\widehat{H}/\partial\bar{\beta}_k=\bar{\beta}_k-\sum_{i=1}^{n}\alpha_{ik}\bar{x}_i=0,\forall k; \tag{18}$$

$$\partial\widehat{H}/\partial\xi_{ik}=C(g'(\xi_{ik}))+\alpha_{ik}=$$
$$C\Big(-c_{ik}+{\rm e}^{\xi_{ik}}/\sum_{k'}^{K}{\rm e}^{\xi_{ik'}}\Big)+\alpha_{ik}=0,\forall i,k. \tag{19}$$

由式 (18) 推导出 $\bar{\beta}_k=\sum_{i=1}^{n}\alpha_{ik}\bar{x}_i,\forall k$. 式 (19) 隐含 $\sum_{k=1}^{K}\alpha_{ik}=0$, 将之应用于式 (17) 可得

$$\alpha_0=0,\sum_{i=1}^{n}\alpha_{ik}=0,\forall k. \tag{20}$$

由式 (19) 还可推导出

$$\xi_{ik}=\log\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)-\frac{1}{K}\sum_{k'=1}^{K}\log\Big(c_{ik'}-\frac{\alpha_{ik'}}{C}\Big), \tag{21}$$

$$g'(\xi_{ik})=-\alpha_{ik}/C. \tag{22}$$

设 $G(\delta)=\delta\xi_{ik}-g(\xi_{ik})$, 其中 $\delta=-\alpha_{ik}/C$. 对 G 进行微分可得

$$\frac{\partial G}{\partial\delta}=\xi_{ik}+\delta\frac{{\rm d}\xi_{ik}}{{\rm d}\delta}-g'(\xi_{ik})\frac{{\rm d}\xi_{ik}}{{\rm d}\delta}=\xi_{ik}. \tag{23}$$

于是 G 可由式 (21) 通过积分得到, 即

$$G\Big(-\frac{\alpha_{ik}}{C}\Big)=$$
$$\frac{K-1}{K}\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)\log\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)+ct. \tag{24}$$

其中: ct 为常量, G 为目标泛函的一部分. 将 Wolfe 对偶理论应用于目标泛函 (13), 并考虑最优化条件 (17) ~(19), 经过简化可得目标泛函的对偶形式为

$$\min D=C\sum_{i=1}^{n}\sum_{k=1}^{K}G\Big(-\frac{\alpha_{ik}}{C}\Big)+\frac{1}{2}\sum_{k=1}^{K}||\bar{\beta}_k||^2, \tag{25}$$

$$\text{s.t.}\sum_{k=1}^{K}\alpha_{ik}=0,\ \ \forall i;\sum_{i=1}^{n}\alpha_{ik}=0,\ \ \forall k. \tag{26}$$

4.2 最优化条件

式 (25) 含有两个约束条件, 要应用最小序列优化进行目标最小化, 首先将其中一个约束条件集成到目标泛函中, 即

$$\min\tilde{D}=C\sum_{i=1}^{n}\sum_{k=1}^{K}G\Big(-\frac{\alpha_{ik}}{C}\Big)+\frac{1}{2}\sum_{k=1}^{K-1}||\bar{\beta}_k||^2+$$
$$\frac{1}{2}\Big|\Big|-\sum_{i=1}^{n}\Big(\sum_{k'=1}^{K-1}\alpha_{ik}\Big)\bar{x}_i\Big|\Big|^2; \tag{27}$$

$$\text{s.t.}\ \sum_{i=1}^{n}\alpha_{ik}=0,\forall k. \tag{28}$$

则对偶目标泛函的拉格朗日形式为

$$\min\widehat{D}=C\sum_{i=1}^{n}\sum_{k=1}^{K}G\Big(-\frac{\alpha_{ik}}{C}\Big)+\frac{1}{2}\sum_{k=1}^{K-1}||\bar{\beta}_k||^2+$$
$$\frac{1}{2}\Big|\Big|-\sum_{i=1}^{n}\Big(\sum_{k'=1}^{K-1}\alpha_{ik}\Big)\bar{x}_i\Big|\Big|^2\sum_{k=1}^{K}\Big(\beta_k\sum_{i=1}^{n}\alpha_{ik}\Big). \tag{29}$$

最小序列优化的最优化条件即是上式对偶函数中 α_{ik} 参数的更改停止条件, 将 $\widehat{D}$ 对 α_{ik} 进行微分可得

$$\frac{\partial\widehat{D}}{\partial\alpha_{ik}}=(\bar{\beta}_k-\bar{\beta}_K)^{\rm T}\bar{x}_i-\Big[\log\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)-$$
$$\log\Big(1-\sum_{k=1}^{K-1}\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)\Big)\Big]-\beta_k. \tag{30}$$

其中 α_{ik} 满足如下条件:

$$\begin{cases}0<\alpha_{ik}<C,c_{ik}=1;\\-C<\alpha_{ik}<0,c_{ik}=0;\end{cases}$$

且

$$0<\sum_{i=1}^{n}\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)<1. \tag{31}$$

设

$$H_{ik}=$$
$$\sum_{t=1}^{n}\alpha_{tk}K(\bar{x}_t,\bar{x}_i)+\sum_{t=1}^{n}\Big(\sum_{k'=1}^{K-1}\alpha_{tk'}K(\bar{x}_t,\bar{x}_i)\Big)-$$
$$\Big[\log\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)+\log\Big(1-\sum_{k=1}^{K-1}\Big(c_{ik}-\frac{\alpha_{ik}}{C}\Big)\Big)\Big], \tag{32}$$

则式 (30) 可表达为 $H_{ik}-\beta_k=0$. 定义

$$\mathrm{upper}(k)=\arg\max_i H_{ik}, \tag{33}$$

$$\mathrm{lower}(k)=\arg\min_i H_{ik}, k=1,2,\cdots,K-1, \tag{34}$$

则多元核Logistic回归模型对偶训练的最优化条件为

$$H_{\mathrm{upper}(k),k}=H_{\mathrm{lower}(k),k}=\beta_k,$$
$$k=1,2,\cdots,K-1. \tag{35}$$

4.3 最小序列优化

基于以上推导的对偶目标泛函及其最优化条件, 多元核Logistic模型的最小序列优化方法与支持向量机类似, 基本操作都是 α 参数的正确初始化以及每次迭代中 $\alpha_{\mathrm{upper}(k)}$ 与 $\alpha_{\mathrm{lower}(k)}$ 的更新, 具体算法流程如下:

Step1: 依照条件 (26) 和 (31), 给定初始化 α 向量, 迭代 $\mathrm{Iter}=1$.

Step2: 如果存在不同索引对 (i,i'), 使得 $H_{i,k}\neq H_{i',k}$, 则依照条件 (33) 和 (34) 选出相应的 $\mathrm{upper}(k)$ 以及 $\mathrm{lower}(k)$.

Step3: 将 $\alpha_{\mathrm{upper}(k),k}$, $\alpha_{\mathrm{lower}(k),k}$ 作如下更新:

$$\alpha_{\mathrm{upper}(k),k}{}^{\mathrm{Iter}+1}=\alpha_{\mathrm{upper}(k),k}{}^{\mathrm{Iter}}+t^*; \tag{36}$$

$$\alpha_{\mathrm{lower}(k),k}{}^{\mathrm{Iter}+1}=\alpha_{\mathrm{lower}(k),k}{}^{\mathrm{Iter}}-t^*; \tag{37}$$

$$\alpha_{i,k}{}^{\mathrm{Iter}+1}=\alpha_{i,k}{}^{\mathrm{Iter}}\text{for other } i,k. \tag{38}$$

Step4: 将 $\alpha^{\mathrm{Iter}+1}$ 代入式 (32) 计算新的 H_{ik}, 并选出新的 $\mathrm{upper}(k)$ 以及 $\mathrm{lower}(k)$.

Step5: 如果对于不同的 $k\in\{1,2,\cdots,K-1\}$, 任意的 (i,i') 索引值对总是满足 $H_{ik}=H_{i'k}$, 则迭代停止; 否则转入 Step2, 直到满足停止条件 (35).

具体实现过程中有以下几点需要注意:

1) 条件 $H_{ik}=H_{i'k}=\beta_k$ 不能严格满足, 因此必须给出相对宽松的条件, 本文以 $H_{\max}-H_{\min}\leqslant 0.001$ 替代.

2) 迭代流程 Step3 中的 t^* 值选择为单变量最优化问题, 可通过牛顿-拉斐逊方法获取, t 变量域必须满足使更改后的 α 依然符合条件 (26) 和 (31). 一般选择 $t_0=0$ 开始进行迭代直到最优, 迭代表达式为

$$t_{\mathrm{new}}=t_{\mathrm{old}}-(\Psi''(t))^{-1}\Psi'(t), \tag{39}$$

其中

$$\Psi(t)=\tilde{D}(\bar{\alpha})=$$
$$C\sum_{i=1}^{n}\sum_{k=1}^{K}G\Big(\frac{\alpha_{ik}}{C}\Big)+\frac{1}{2}\sum_{k=1}^{K-1}||\bar{\beta}_k||^2+\frac{1}{2}||\bar{\beta}_K||^2.$$

则有

$$\Psi'(t)=H_{ik}-H_{i'k}, \tag{40}$$

$$\Psi''(t)=$$
$$2[(K\bar{x}_i,\bar{x}_i)-2K(\bar{x}_i,\bar{x}_{i'})+K(\bar{x}_{i'},\bar{x}_{i'})]+$$
$$(Cc_{ik}-\alpha_{ik}-t)^{-1}+(Cc_{i'k}-\alpha_{i'k}+t)^{-1}+$$
$$\Big(C-\sum_{k=1}^{K-1}(Cc_{ik}-\alpha_{ik})+t\Big)^{-1}+$$
$$\Big(C-\sum_{k=1}^{K-1}(Cc_{i'k}-\alpha_{i'k})-t\Big)^{-1}. \tag{41}$$

3) H_{ik} 函数是迭代优化的对象, 且每次 α 更新过程都需要用到, 因此必须进行及时的更新保存, 并应用于下一次迭代操作, 设

$$\mathrm{sa}(p,\alpha)=\sum_{k=1}^{K-1}\Big(c_{p,k}-\frac{\alpha_{p,k}}{C}\Big), \tag{42}$$

则有

$$H_{i,k}^{\mathrm{new}}=$$
$$H_{i,k}^{\mathrm{old}}+2t[K(\bar{x}_i,\bar{x}_i)-K(\bar{x}_i,\bar{x}_{i'})]-$$
$$\log\Big(c_{ik}-\frac{\alpha_{ik}^{\mathrm{new}}}{C}\Big)+\log\Big(c_{ik}-\frac{\alpha_{ik}^{\mathrm{old}}}{C}\Big)+$$
$$\log[1-\mathrm{sa}(i,\alpha_{\mathrm{new}})]-\log[1-\mathrm{sa}(i,\alpha_{\mathrm{old}})], \tag{43}$$

$$H_{i',k}^{\mathrm{new}}=$$
$$H_{i',k}^{\mathrm{old}}+2t[K(x_{i'},x_i)-K(x_{i'},x_{i'})]-$$
$$\log\Big(c_{i'k}-\frac{\alpha_{i'k}^{\mathrm{new}}}{C}\Big)+\log\Big(c_{i'k}-\frac{\alpha_{i'k}^{\mathrm{old}}}{C}\Big)+$$
$$\log[1-\mathrm{sa}(i',\alpha_{\mathrm{new}}]-\log[1-\mathrm{sa}(i',\alpha_{\mathrm{old}})], \tag{44}$$

$$H_{i,j}=$$
$$t[K(\bar{x}_i,\bar{x}_i)-K(\bar{x}_i,\bar{x}_{i'})]+$$
$$\log(1-\mathrm{sa}(i,\alpha_{\mathrm{new}}))-\log(1-\mathrm{sa}(i,\alpha_{\mathrm{old}})), \tag{45}$$

$$H_{i',j}=$$
$$t[K(\bar{x}_i,\bar{x}_{i'})-K(\bar{x}_{i'},\bar{x}_{i'})]+$$
$$\log(1-\mathrm{sa}(i',\alpha_{\mathrm{new}}))-\log(1-\mathrm{sa}(i',\alpha_{\mathrm{old}})), \tag{46}$$

$$H_{m,j}=t[K(x_i,x_m)-K(x_{i'},x_m)]. \tag{47}$$

其中: 每次迭代过程中 $i=\mathrm{upper}(k)$, $i'=\mathrm{lower}(k)$, 且 $j\neq k$, $m\neq\mathrm{uppper}(k)$, $m\neq\mathrm{lower}(k)$.

4.4 话者辨别实验

实验对以上所提多元核 Logistic 回归模型进行

说话人辨别应用, 验证其算法的有效性与在说话人辨别应用场合的可行性. 采用自己录制的语料库, 录音总人数为 20, 其中男 12 人, 女 8 人. 数据通过采样频率 8 000 Hz, 量化位数 16 bit, 单声道 A/D 转化获得. 每个人的语音信号通过不同时期录制合成. 每人混合提取不同时期的语音片段总长度 15 s作为训练信号, 不同时期的 20 个长度为 1.5 s的语音片段作为测试信号, 即 20 个训练语音, 400 个测试语音. 语音信号先经高频提升. 去直流等预处理, 再通过声音活性检测, 提取其中有效的语音段, 去除冗余的无声段, 以 30 ms 长度分帧提取 12 维的 MFCC (mel frequency cepstral coefficient) 特征参数, 即 mel 频率倒谱系数, 作为分类参数.

实验都在相同的平台中完成: 操作系统: Windows XP; 计算环境: Matlab 7.1; 台式机电脑: Intel 酷睿双核, 主频 2.6 GHz, 内存 4 G. 首先将多元核 Logistic 回归对偶训练算法与文献 [12] 的 IRRLS 算法进行运算效率对比, 本文算法的每次迭代过程解决了最小序列优化问题, 而[12]直接对原型负对数似然目标泛函进行优化. 因此, 二者停止准则不一致, 直接对比效率显得不甚公平. 对此采取的方法是先用对偶算法进行说话人训练, 记录用时, 根据满足要求的 α 参数计算 β 向量; 然后利用 [12] 的 IRRLS 算法对相同训练样本进行迭代优化, 记录达到近似 β 向量所需时间, 如表 1 所示.

表 1　两种多元核 Logistic 回归训练算法效率对比

训练语音长度/s	样本数×维数	对偶算法用时/s	IRRLS算法用时/s
5	4 982×12	5.86	380.6
10	9 961×12	11.35	916.4
15	14 945×12	23.58	2.8e3

从表 1 可见, 对于不同训练长度的语音信号样本, 对偶算法训练速度都优于 IRRLS 算法, 且训练时间随样本量的增长基本呈线性关系, 可以应用于说话人辨别场合.

核函数在基于核方法的分类场合中起到具有重要作用, 应用不同核函数的多元核 Logistic 回归模型说话人辨别识别率也差异甚大. 在以下将 3 种典型的核函数应用于对偶训练的算法中, 对 10 个说话人进行分类测试. 这 3 种核函数分别为径向基核函数, 三阶多项式核函数和 Cauchy 核函数. 不同函数的参数 σ 取值都通过交叉验证获取. 另外, 线性核函数由于不适应于说话人辨别这类复杂非线性场合, 在此不作比较. 表 2 列出了基于不同核函数的说话人辨别识别率的比较结果.

由表 2 可知, 在几种经典核函数对比测试中, 径向基函数 (RBF) 具有明显的性能优势. 因此, 在说话人辨别实验中, 均采用径向基内核进行非线性分类.

表 2　多元核 Logistic 回归应用不同核函数的说话人辨别对比

核函数	识别率/%	σ
RBF	100	1 000
Poly	86	0.8
Cauchy	92.5	1.9

将多元核 Logistic 回归方法与二元核 Logistic 回归方法以及支持向量机方法进行说话人辨别识别率对比. 其中后两者都是二元分类器, 都采用构建"一对一"多个分类模型进行投票式识别的方法, 支持向量机也采用相同的径向基核函数, σ 取值 1.5. 对比结果如表 3 所示.

表 3　两种多元核 Logistic 回归训练算法效率对比　%

算 法	效 率		
	5 人	10 人	20 人
MultiKLR	0/100=0	2/200=1	9/400=2.25
SVM	0/100=0	2/200=1	10/400=2.5
2_KLR	0/100=0	1/200=0.5	11/400=2.75

由表 3 可知, 核 Logistic 回归方法与经典说话人辨别所采用的支持向量机法在识别率上相似, 但本文所提方法简单直观, 每次识别任务中只需进行一次模型构建即可应用. 而二元核 Logistic 回归方法与支持向量机方法在应用于多分类场合时, 必须构建多个二元分类器以间接实现多分类. 如本例进行 20 人辩别实验时, 采用"一对一"多分类方法的支持向量机与二元核 Logistic 回归必须构建 $C_{20}^2 = 190$ 个不同的二元分类器, 这大大增加了训练模型的复杂度与繁琐性. 相比较而言, 多元方法更加便捷与人性化.

5　结　论

针对说话人辨别的非线性、多元性等特点, 将传统核 Logistic 回归扩展为多元化核 Logistic 回归模型, 并将目标泛函修改为对偶形式以适应支持向量机中的最小优化算法应用. 经验证, 该算法速率快于传统多元核 Logistic 回归训练法, 且训练时间随训练样本数增长基本呈线性关系, 符合说话人辨别场合训练数据量大的情况, 获得与传统支持向量机相似的识别率, 且模型构建方便快捷, 一次成型. 缺点是训练结果不具稀疏性, 在识别应用中的实时性能差, 今后将在这方面进行研究, 在识别率不降低的前提下提高识别实时性.

参考文献(References)

[1] Frederic Bimbot, Jean-Francis Bonastre. A tutorial on text-independent speaker verification[J]. Eurasip J on Applied Signal Processing, 2004, (4): 430-451.

[2] Wan V, Renals S. Speaker verification using sequence discriminant support vector machines[J]. Speech and Audio processing, 2005, 13(2): 203-210.

[3] 李滔, 王俊普, 吴秀清. 基于特征矢量集的核 Logistic 回归[J]. 小型微型计算机系统, 2006, 27(6): 980-985.
(Li T, Wang J P, Wu X Q. Kernel Logistic regression based on feature vector set[J]. J of Chinese Computer System, 2006, 27(6): 980-985.)

[4] Volker Roth. Probabilistic discriminative kernel classiffiers for multi-class problems[C]. Proc of the 23rd DAGM-Symposium on Pattern Recognition. London: Springer-Verlag, 2001: 246-253.

[5] Shevade S K, Keerthi S S. A simple and efficient algorithm for gene selection using sparse Logistic regression[J]. Bioinformatics, 2003, 19(17): 2246-2253.

[6] Rahayu, Purnami S P, Embong S W. Applying kernel Logistic regression in data mining to classify credit risk[C]. Int Symposium on Information Technology. Kuala Lumpur, 2008, 2: 1-6.

[7] Birkenesm O, Matsui T, Tanabe K. Isolated-word recognition with penalized Logistic regression machines [C]. Acoustics, Speech and Signal Processing. Toulouse, 2006: 405-408.

[8] Katz M, Schaffoner M, Andelic E, et al. Sparse kernel Logistic regression using incremental feature selection for text-independent speaker identification[C]. Speaker and Language Recognition workshop. San Juan, 2006: 1-6.

[9] Katz M, Schaffner M, Andelic E, et al. Sparse kernel Logistic regression for phoneme classification[C]. Proc of 10th Int Conf on Speech and Computer. Patras, 2005, 2: 523-526.

[10] Roth V, Stenhage V. Nonlinear discrminant analysis using kernel functions[C]. Advances in Neural Information Processing Systems. Cambridge, 1999: 558-574.

[11] Karsmakers P, Pelckmans K, Suykens J A K. Multi-class kernel Logistic regression: A fixed-size implementation [C]. The 2007 IEEE Int Joint Conf on Neural Networks. Orlando, 2007: 1756-1761.

[12] Keerthi S S, Duan K B, Shevade S K, et al. A fast dual algorithm for kernel Logistic regression[J]. Machine Learning, 2005, 61(1-3): 151-165.

(上接第1434页)

[5] Jabr R A. Robust self-scheduling under price uncertainty using conditional value-at-risk[J].IEEE Trans Power Systems, 2005, 20(4): 1852-1858.

[6] Zhou S S, Fukushima M. Worst-case conditional value-at-risk with application to robust portfolio management[J]. Operations Research, 2009, 57(5): 1155-1168.

[7] 徐玖平, 李军. 多目标决策的理论与方法[M]. 北京: 清华大学出版社, 2005.
(Xu J P, Li J. Multiple objective decision making theroy and method[M]. Beijing: Tsinghua University Press, 2005.)

[8] Ogryczak W. Multiple criteria linear programming model for portfolio selection[J]. Annals of Operations Research, 2000, (97): 143-162.

[9] Ogryczak W. Multicriteria models for fair resource allocation[J]. Control and Cybernetics, 2007, (36): 303-332.

[10] Mansini R, Ogryczak W, Speranza M G. Conditional value at risk and related linear programming models for portfolio optimization[J]. Annals of Operations Research, 2007, (152): 227-256.

[11] 刘敏, 吴复立. 电力市场环境下发电商电能分配策略研究[J]. 中国电机工程学报, 2008, 28(25): 111-117.
(Liu M , Wu F. Trading strategy of generation companies in electricity market[J]. Proc of the CSEE, 2008, 28(25): 111-117.)

[12] Su J. Analytical assessment of generation asset in restructured electricity industry[D]. Hong Kong: University of Hong Kong, 2006.