

文章编号: 1001-0920(2013)05-0683-05

基于流形学习的局部保持 PCA 算法在故障检测中的应用

王 健^{1,2}, 冯 健¹, 韩志艳²

(1. 东北大学 信息科学与工程学院, 沈阳 110819; 2. 渤海大学 工学院, 辽宁 锦州 121000)

摘 要: 提出一种新的基于流形学习的数据降维及特征提取方法: 局部保持 PCA 算法 (LPPCA). 通过在 PCA 的优化目标中融入流形学习的思想, 不仅使投影得到的低维空间和原始样本空间具有相似的全局结构, 并且保持了相似的局部近邻结构, 克服了传统 PCA 方法只关注全局结构特征而忽略局部流形特征的缺陷, 同时给出了 LPPCA 在故障检测中的应用方法. S-Curve 和 Swiss-roll 曲面数值仿真和 TE 过程仿真验证了算法的有效性和优越性.

关键词: 主元分析; 局部保持; 故障检测; 流形学习

中图分类号: TP273

文献标志码: A

Locally preserving PCA method based on manifold learning and its application in fault detection

WANG Jian^{1,2}, FENG Jian¹, HAN Zhi-yan²

(1. College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 2. College of Engineering, Bohai University, Jinzhou 121000, China. Correspondent: FENG Jian, E-mail: fjneu@163.com)

Abstract: A novel dimensionality reduction and feature extraction method based on manifold learning, locally preserving principal component analysis(LPPCA) is proposed. In order to overcome the defects that the traditional PCA can only keep the structure in global and can not maintain the manifold structure in local, the idea of locality preserving is incorporated into the optimization goals of the PCA. The fault detection based on LPPCA is researched. The validity and superiority of the LPPCA are verified by the S-Curve numerical simulation, Swiss-roll surface numerical simulation and TE process simulation.

Key words: principal component analysis(PCA); locality preserving projections(LPP); fault detection; manifold learning

0 引 言

在过去的 10 年里, 多元统计过程监控 (MSPM) 方法受到了学术界和工程界的广泛关注和研究^[1-3]. MSPM 包括: 主元分析 (PCA), 部分最小二乘 (PLS) 和独立主元分析 (ICA) 等. 其中 PCA 是目前广泛应用于过程监控的 MSPM 方法, 是一种将多个相关变量转化为少数几个相互独立变量的有效分析方法^[4], 它不依赖于过程机理, 只需通过过程数据的信息进行统计建模, 即可实现以较少维数的数据描述监控对象的主要特征. 然而, PCA 方法只能发现过程数据的全局欧几里德结构, 忽略了数据的局部结构特征. 最近, 大量研究表明^[5-7], 流形学习能够有效地发现隐含在高维数据集中的低维特征, 对局部结构特征具有较好的提取

能力. 多种流形学习算法也被提出, 如: 局部线性嵌入 (LLE)、保距特征映射 (Isomap)、拉普拉斯特征映射 (LE) 等. 其中, LE 的线性化映射——局部保持投影法 (LPP) 在继承 LE 算法能够保持局部流形特征的基础上, 实现了线性计算, 得到广泛的应用.

本文将 LPP 算法的思想融入 PCA 的目标函数中, 使得在原始样本空间投影到低维空间的过程中, 不仅实现了整体方差最大化, 而且保持了局部近邻结构不变, 有利于全局和局部特征的全面提取. 因此, 把新提出的算法命名为局部保持 PCA 算法 (LPPCA).

1 LPPCA 算法

作为特征提取和降维的经典方法, PCA 通过把数据投影到方差最大的方向, 保留了数据的全局结

收稿日期: 2012-05-02; 修回日期: 2012-07-17.

基金项目: 国家自然科学基金重点项目(61273164, 61034005); 国家高技术研究发展计划项目(2012AA040104); 中央高校基本科研业务费项目(N100104102, N120504002).

作者简介: 王健(1982-), 男, 讲师, 博士生, 从事故障诊断、数据挖掘等研究; 冯健(1971-), 男, 教授, 博士生导师, 从事故障诊断、信号处理、智能控制等研究.

构特征. 但是, PCA方法建立在欧氏空间的线性理论基础上, 在实际环境中, 此条件常常难以满足, 这就意味着在某些复杂情况下, 仅仅依靠PCA很难有效提取信息的全面特征. LPP方法的出现对解决这一问题提供了有力的途径, LPP通过构造近邻图来探索数据的内在的几何特征和流形结构, 不仅弥补了传统PCA忽略了数据局部流形结构的缺陷, 而且由于LPP是利用局部结构来描述全局特征, 在一定程度上, 解决了PCA线性方法的局限性.

结合PCA和LPP算法的优点, 本文提出一种新的数据降维和特征提取方法——LPPCA. 假设样本集 $X = [x_1, x_2, \dots, x_n] \in R^M$ 为原始样本空间的一组采样, LPPCA的基本思想是寻找一个投影矩阵 $W \in R^{M \times d}$, 使得 X 通过投影 $y_i = W^T x_i (i = 1, 2, \dots, n)$ 得到的低维映射 $Y = [y_1, y_2, \dots, y_n] \in R^d (d < M)$ 不仅和原始样本空间具有相似的局部近邻结构, 而且保留了原始样本空间的大部分方差信息, 实现了全局和局部特征的全面提取. 因此, LPPCA的优化目标可以分解为全局和局部两个目标函数, 全局目标函数使映射得到的低维空间可以保留原始样本空间的大部分方差信息, 实现了全局特征提取; 局部目标函数使映射得到的低维空间和原始样本空间具有相似的局部近似结构, 实现了局部特征提取.

1.1 全局目标函数描述

假设已知样本集 $X = [x_1, x_2, \dots, x_n]$, $J(W)_{\text{global}}$ 的目标是寻找 d 个投影向量 w_1, w_2, \dots, w_d , 构成投影矩阵 W , 使得通过投影映射 $y_i = W^T x_i (i = 1, 2, \dots, n)$ 得到的低维空间保留原始样本空间的大部分方差信息, 即

$$\begin{aligned} J(W)_{\text{global}} &= \max_W \sum_{i=1}^n \|y_i - \bar{y}\|^2 = \\ &= \max_W \sum_{i=1}^n W^T (x_i - \bar{x})(x_i - \bar{x})^T W = \\ &= \max_W W^T C W. \end{aligned} \quad (1)$$

其中

$$\begin{aligned} \bar{y} &= \left(\sum_{i=1}^n y_i \right) / n, \quad \bar{x} = \left(\sum_{i=1}^n x_i \right) / n, \\ C &= ((x_i - \bar{x})(x_i - \bar{x})^T) / n, \\ W &= [w_1, w_2, \dots, w_d]. \end{aligned}$$

可以看出, LPPCA的全局目标函数与PCA的目标函数是一致的, 都是通过线性空间变换将高维数据空间投影到低维空间. 因为低维空间保留了原始样本空间的大部分方差信息, 所以实现了全局结构的特征提取. 但是, 全局目标函数中没有考虑样本点的局部几何关系, 在低维空间里, 样本点之间的局部几何关

系有可能被打乱, 导致过程信息的丢失.

1.2 局部目标函数描述

假设已知样本集 $X = [x_1, x_2, \dots, x_n]$, $J(W)_{\text{local}}$ 的目标是寻找 d 个投影向量 w_1, w_2, \dots, w_d , 构成投影矩阵 W , 使得通过投影映射 $y_i = W^T x_i (i = 1, 2, \dots, n)$ 得到的低维空间和原始样本空间具有相似的局部近邻结构. 最佳映射可通过解决下面的最小化问题得到:

$$\begin{aligned} J(W)_{\text{local}} &= \\ \min_W \frac{1}{2} \sum_{i,j=1}^n \|y_i - y_j\|^2 S_{ij} &= \\ \min_W \frac{1}{2} \sum_{i,j=1}^n (W^T x_i - W^T x_j)(W^T x_i - W^T x_j)^T S_{ij} &= \\ \min_W \left(\sum_{i,j=1}^n W^T x_i S_{ij} x_j^T W - \sum_{i,j=1}^n W^T x_i S_{ij} x_j^T W \right) &= \\ \min_W \left(\sum_{i=1}^n W^T x_i D_{ii} x_i^T W - W^T X S X^T W \right) &= \\ \min_W \left(W^T X D X^T W - W^T X S X^T W \right) &= \\ \min_W W^T X (D - S) X^T W &= \\ \min_W W^T X L X^T W. \end{aligned} \quad (2)$$

其中: S_{ij} 为 x_i 和 x_j 之间的近邻关系, 一般取

$$S_{ij} = \begin{cases} e^{-\|x_i - x_j\|^2 / t}, & x_i \text{ 和 } x_j \text{ 是近邻;} \\ 0, & \text{否则.} \end{cases}$$

s.t. $0 < i < n, 0 < j < n.$ (3)

S 为权重矩阵, D 为对角阵, $D_{ii} = \sum_j S_{ij}$; $L = D - S$ 是Laplacian矩阵; $W = [w_1, w_2, \dots, w_d]$ 是投影矩阵.

可以看出, LPPCA的局部目标函数和LPP的目标函数是一致的, 都是通过使在原始样本空间中相邻的点映射到低维空间中也相邻, 完成对样本点局部流形结构特征的提取. 但是, 局部目标函数没有显式地考虑样本点的全局特征, 仅仅利用局部结构来描述全局特征, 有可能导致在低维空间中全局特征的扭曲.

1.3 LPPCA算法描述

为结合PCA和LPP算法的优点, 本文将LPP算法的思想融入PCA的目标函数中, 使得LPPCA投影得到的低维空间能够获得更全面的特征信息. LPPCA的全局目标函数定义为

$$J(W)_{\text{global}} = \max_W W^T C W, \quad (4)$$

LPPCA的局部目标函数定义为

$$\begin{aligned} J(W)_{\text{local}} &= \min_W W^T X L X^T W = \\ \min_W (W^T X D X^T W - W^T X S X^T W). \end{aligned} \quad (5)$$

当 $W^T X D X^T W = 1$ 时, 得到^[8-9]

$$J(W)_{local} = 1 - \min_W W^T X S X^T W = \max_W W^T X S X^T W. \quad (6)$$

综合考虑全局和局部目标函数, LPPCA 算法的目标函数定义为

$$J(W)_{LPPCA} = \max_W (W^T C W + \lambda W^T X S X^T W) = \max_W (W^T (C + \lambda X S X^T) W) = \max_W W^T \tilde{S} W, \quad (7)$$

其中 λ 为平滑参数, 用以控制全局特征和局部特征的提取比例, λ 越小越侧重于全局特征的提取, λ 越大越侧重于局部特征的提取.

式(7)的求解可通过求取 \tilde{S} 最大的 d 个特征值 a_1, a_2, \dots, a_d 所对应的特征向量 w_1, w_2, \dots, w_d 完成, 即投影矩阵 $W = [w_1, w_2, \dots, w_d]$, 低维投影通过映射 $Y = W^T X$ 得到.

2 LPPCA 算法在故障检测中的应用

LPPCA 故障检测方法在离线建模阶段对正常工况训练数据实施 LPPCA, 得到 LPPCA 投影矩阵 W 及训练数据在 LPPCA 模型空间中的投影. 在线监测阶段对在 LPPCA 模型空间中的投影得到的低维空间和对应的残差空间分别实施 T^2 和 SPE 监测来实现故障检测. 具体步骤如下:

Step 1: 收集正常工况下的训练数据, 并归一化, 得到原始样本集 X ;

Step 2: 对样本集 X 实施 LPPCA, 得到投影矩阵 W ;

Step 3: 计算样本集 X 在 LPPCA 子空间中的低维投影 Y 和对应的残差空间;

Step 4: 计算与样本集 X 相关联的 T^2 和 SPE 统计量, 得到 T^2 和 SPE 统计量控制上限;

Step 5: 获得当前监控数据并归一化;

Step 6: 使用 Step 2 获得的投影矩阵 W 把监控样本投影到低维空间, 并计算样本的 T^2 和 SPE 统计量;

Step 7: 观察统计量是否超过各自的统计限.

3 仿真研究

3.1 数值仿真

S-Curve 曲面和 Swiss-roll 曲面是两个经常使用的人造数据, 本文在 S-Curve 和 Swiss-roll 曲面上随机采样 2000 个数据点, 详细比较 PCA、LPP 和 LPPCA 算法的性能. 实验中选取 LPP 和 LPPCA 算法的近邻参数 $k = 10$, 并按式(3)构造权重矩阵 S . LPPCA 中 $\lambda = 0.5$. 图1和图2分别显示了 PCA、LPP 和 LPPCA 三种

算法在 S-Curve 曲面和 Swiss-roll 曲面上的降维效果. 可以看出 PCA 在 S-Curve 和 Swiss-roll 上的降维效果并不好, 特别是对数据的分布边界保持较模糊; LPPCA 在 3 种方法中对数据的分布边界保持最清晰, 故障检测在离线建模阶段的目标实质是要得到过程正常工况数据波动的边界, 以此作为判断在线数据是否为故障的依据, 在这个意义上 LPPCA 是一种理想的降维方法.

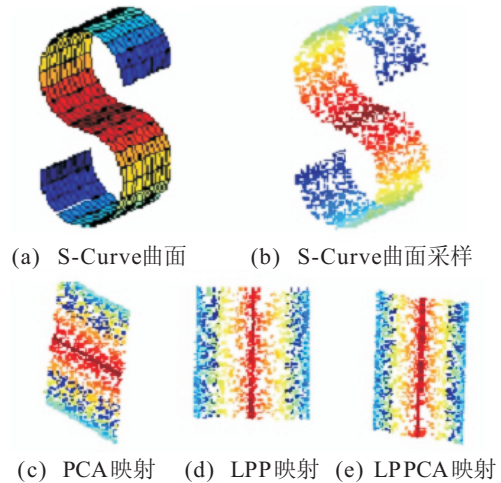


图1 S-Curve 曲面实验

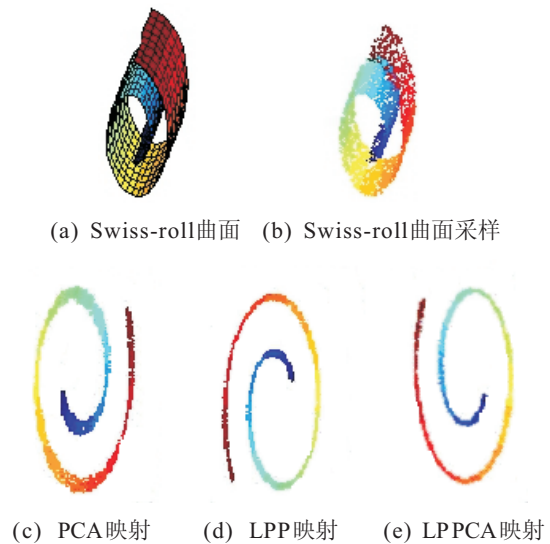
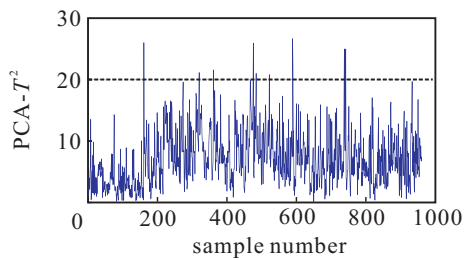


图2 Swiss-roll 曲面实验

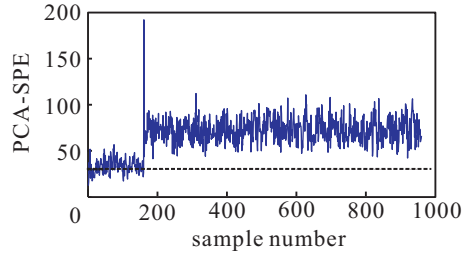
3.2 TE 过程仿真

TE 过程是一个公认的比较各种控制和监控方案的平台, 共设有 52 个过程变量, 6 种操作模式. 过程的详细描述、工艺流程图以及其故障形式的具体介绍和使用见文献[10]. 本文使用的 TE 过程数据下载自 <http://brahms.scs.uiuc.edu>. 训练数据集包含 500 个在过程正常状态下采集的数据点. 每个测试数据集中有一种故障模式, 包含 960 个数据点, 故障从第 161 个数据点开始引入. 过程包括 12 个操作变量和 41 个测量变量, 选择了 52 个变量用于监测. 建立的 PCA 模

型选取的主元数目为 8 个, 累计方差贡献率达到 95%. 同样, 建立 LPPCA 故障检测模型时, 选取低维的维度也为 8, 所取得置信限为 90%; 同时, LPPCA 算法的近邻选取参数 $k = 10$, 按式 (3) 构造权重矩阵 S , LPPCA 中 $\lambda = 0.5$. 为检验 LPPCA 方法在故障诊断中的效果, 选取故障 4 和故障 9 进行验证性仿真研究, 仿真结果如图 3~图 6 所示. 故障 4 发生时的主要表现是反应器冷却水流速的阶跃变化, 反应器温度突然升高. 从图 3 可以看出, PCA 的 SPE 统计量可以对故障 4 实施有效检测, 但 T^2 统计量对故障 4 的反应不敏感; 图 4 的结果显示, LPPCA 的 SPE 和 T^2 统计量均可以对故障 4 实施有效检测. 这说明 LPPCA 不仅继承了 PCA 的检测效果, 而且由于融入了局部几何结构特征, 对故障特征的提取更加全面. 故障 9 发生时的主

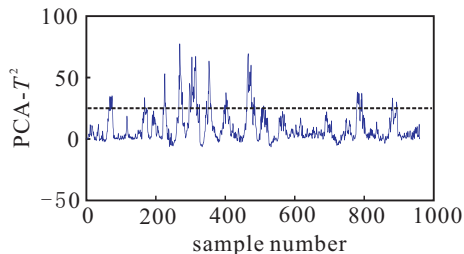


(a) T^2 统计量

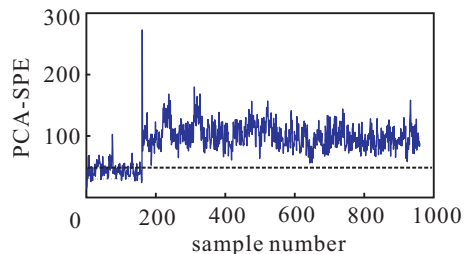


(b) SPE 统计量

图 3 PCA 对故障 4 的检测

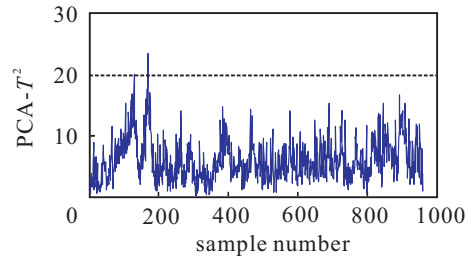


(a) T^2 统计量

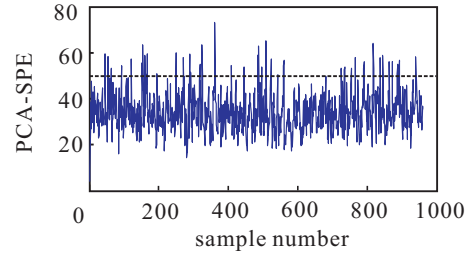


(b) SPE 统计量

图 4 LPPCA 对故障 4 的检测

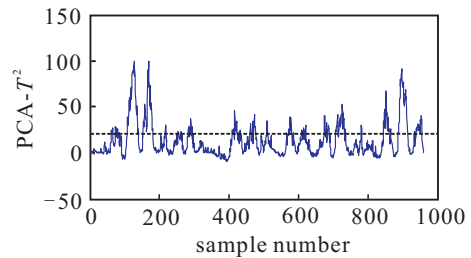


(a) T^2 统计量

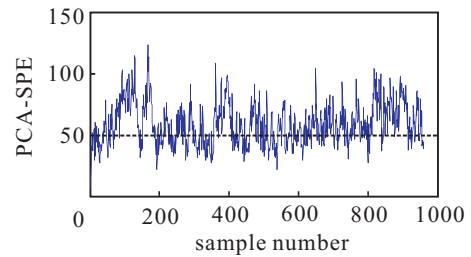


(b) SPE 统计量

图 5 PCA 对故障 9 的检测



(a) T^2 统计量



(b) SPE 统计量

图 6 LPPCA 对故障 9 的检测

要表现是进料 D 的温度变化, 此故障的采样数据和正常工况下的数据差别不大, 因此难以检测. 从图 5 可看出, PCA 算法对故障 9 这种存在严重非线性的工业过程故障检测的效果很差, 在故障发生时, SPE 和 T^2 统计量对故障的检测失效; 图 6 中 LPPCA 的 SPE 统计量大幅超过控制限, T^2 统计量也对故障的检测有效, 表明 LPPCA 对非线性故障检测有效, 通过融入局部几何结构特征, 提高了对非线性故障特征的提取能力.

4 结论

本文提出了一种新的数据降维及特征提取方法 LPPCA, 并将其应用于故障检测中. 该方法不仅实现了全局特征的保持, 而且兼顾了局部特征的提取, 并且可以通过适当调整平滑参数 λ 来改变整体特征和

局部特征的保持比例. 通过 S-Curve 曲面和 Swiss-roll 曲面数值仿真, 验证了该算法的降维效果; 通过 TE 过程, 验证了该算法在故障检测中的有效性和优越性. 平滑参数 λ 对 LPPCA 算法的整体性能具有重要影响, 如何定性地确定 λ 的取值范围是将来工作的研究方向之一.

参考文献(References)

- [1] 周东华, 李钢, 李元. 数据驱动的工业过程故障诊断与预测技术——基于 PCA 与 PLS 的方法[M]. 北京: 科学出版社, 2011.
(Zhou D H, Li G, Li Y. Data-driven fault diagnosis of industrial processes and forecasting techniques—Based on PCA and PLS methods[M]. Beijing: Science Press, 2011.)
- [2] 李晗, 萧德云. 基于数据驱动的故障诊断方法综述[J]. 控制与决策, 2011, 26(1): 1-9.
(Li H, Xiao D Y. Survey on data driven fault diagnosis methods[J]. Control and Decision, 2011, 26(1): 1-9.)
- [3] Hwang I, Kim S, Kim Y, et al. A survey of fault detection, isolation, and reconfiguration methods[J]. IEEE Trans on Control Systems Technology, 2010, 18(3): 636-653.
- [4] 肖应旺, 徐保国. 改进 PCA 在发酵过程监测与故障诊断中的应用[J]. 控制与决策, 2005, 20(5): 571-574.
(Xiao Y W, Xu B G. Application of improved PCA to fermentation process monitoring and fault diagnosis[J]. Control and Decision, 2005, 20(5): 571-574.)
- [5] 张沐光, 宋执环. LPMVP 算法及其在故障诊断中的应用[J]. 自动化学报, 2009, 35(6): 766-772.
(Zhang M G, Song Z H. LPMVP algorithm and its application to fault detection[J]. Acta Automatica Sinica, 2009, 35(6): 766-772.)
- [6] Shao J D, Rong G. Nonlinear process monitoring based on maximum variance unfolding projections[J]. Expert Systems with Applications, 2009, 36(8): 11332-11340.
- [7] Hu K, Yuan J. Batch process monitoring with tensor factorization[J]. J of Process Control, 2009, 19(2): 288-296.
- [8] He X F, Yan S C. Face recognition using laplacianfaces[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(3): 328-340.
- [9] Deng C, He X F. Locally consistent concept factorization for document clustering[J]. IEEE Trans on Knowledge and Data Engineering, 2011, 23(6): 902-913.
- [10] Cho J H, Lee J M, Choi S W, et al. Fault identification for processes monitoring using kernel principal component analysis[J]. Chemical Engineering Science, 2005, 60(1): 279-288.

(上接第676页)

- [5] Hastie T, Tibshirani R. Classification by pairwise coupling[J]. The Annals of Statistics, 1998, 26(1): 451-471.
- [6] Wu Ting-Fan, Lin Chih-Jen, Ruby C Weng. Probability estimates for multi-class classification by pairwise coupling[J]. J of Machine Learning Research, 2004, 5(8): 975-1005.
- [7] Dietterich T G, Bakiri G. Solving multiclass learning problems via error-correcting output codes[J]. J of Artificial Intelligence Research, 1994, 2(1): 263-286.
- [8] Dietterich T G, Bakiri G. Error-correcting output codes: A general method for improving multiclass inductive learning programs[C]. Proc of the 9th National Conf on Artificial Intelligence(AAAI-91). Menlo Park: AAAI Press, 1991:572-577.
- [9] Pujol O, Radeva P, Vitria J. Discriminate ECOC: A heuristic method for application dependent design of eError correcting output codes[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2006, 28(6): 1001-1007.
- [10] Crammer K Y, Singer on the learnability and design of output codes for multiclass problems[J]. Machine Learning, 2002, 47(2/3): 201-233.