

基于相同稀疏模式的稀疏主成分分析算法

邵剑飞¹, 浦蓉¹, 黄伟², 季建杰¹, 郭鹏¹

(1. 昆明理工大学信息工程与自动化学院, 昆明 650500; 2. 云南电网有限责任公司红河供电局, 红河 661100)

摘要: 稀疏主成分分析是一种用于降维和特征选择的无监督方法。由于计算多个主成分时主载荷向量间不具有相同的稀疏模式, 导致难以从原始特征空间中确定出对主成分贡献最大的小部分变量, 为解决此问题, 提出一种自适应稀疏主成分分析(Adaptive sparse principal component analysis, ASPCA)算法。首先使用组套索模型, 通过在载荷向量上施加块稀疏约束得出自适应稀疏主成分分析公式, 随后对稀疏矩阵的不同列使用不同的调整参数获得自适应惩罚, 最后运用块坐标下降法对自适应稀疏主成分分析公式进行两阶段优化, 从而找到稀疏载荷矩阵和正交矩阵, 实现降维的最优化。对稀疏主成分分析(Sparse principal component analysis, SPCA)算法、结构化且稀疏的主成分分析(Structured and sparse principal component analysis, SSPCA)算法和 ASPCA 算法进行仿真比较, 结果表明 ASPCA 算法的降维性能更优, 能提取更有价值的特征, 从而显著提高了分类模型的平均分类准确率。

关键词: 稀疏主成分分析; 稀疏模式; 主载荷向量; 调整参数; 块坐标下降法; 降维

中图分类号: TP391.41 **文献标志码:** A

Sparse Principal Component Analysis Algorithm Based on Same Sparse Pattern

SHAO Jianfei¹, PU Rong¹, HUANG Wei², JI Jianjie¹, GUO Peng¹

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;
2. Honghe Power Supply Bureau, Yunnan Power Grid Co., Ltd., Honghe 661100, China)

Abstract: Sparse principal component analysis is an unsupervised method for dimensionality reduction and feature selection. An adaptive sparse principal component analysis (ASPCA) algorithm is proposed, because the principal load vectors do not have the same sparse pattern when calculating multiple principal components, and it is difficult to determine a small number of the variables that contribute the most to the principal components from the original feature space. Firstly, the group lasso model is used, and the ASPCA formula is obtained by applying block sparse constraints on the load vector. Subsequently, different adjustment parameters are used for different columns of the sparse matrix to obtain adaptive penalty. Finally, the block-coordinate descent method is used to optimize the adaptive sparse principal component analysis formula in two stages, so as to find the sparse load matrix and the orthogonal matrix and achieve the optimization of dimensionality reduction. The comparison results of the sparse principal component analysis (SPCA) algorithm, the structured and sparse principal component analysis (SSPCA) algorithm and the ASPCA algorithm show that the ASPCA algorithm has better dimensionality reduction performance and can extract more valuable features, thereby effectively improving the average

classification accuracy of the classification model.

Key words: sparse principal component analysis; sparse mode; principal load vector; adjustment parameters; block coordinate descent; dimensionality reduction

引 言

主成分分析(Principal component analysis, PCA)是数据分析中一种流行的特征提取和降维技术^[1-2],被广泛应用于图像和视频分类^[3]、图像去噪^[4]以及面部识别^[5]等应用数据分析^[6]中。但是,传统PCA的主成分是原始变量的线性组合并且载荷通常为非零,因此稠密的主载荷向量导致主成分缺乏可解释性。而稀疏主成分分析方法是计算具有少量非零项的主载荷向量实现特征选择,同时使模型具备更好的可解释性^[7]。Zou等^[8]提出了稀疏主成分分析算法(Sparse principal component analysis, SPCA)。SPCA算法首先证明PCA可以表述为回归型优化问题,然后通过对回归系数施加套索(弹性网)约束来获得稀疏载荷,但该算法不能保证在所有主载荷向量上都保留稀疏模式,即非零变量的位置在不同的主载荷向量上是不同的。Jenatton等^[9]提出了一种结构化且稀疏的主成分分析算法(Structured and sparse principal component analysis, SSPCA)。SSPCA算法定义了范数的组与由此产生的非零模式之间的关系,提供了从组到模式来回的向前和向后算法,设计适合以非零模式表示的特定先验知识的规范。该算法对主载荷向量的稀疏模式施加约束,但是由于SSPCA算法需要进行通用优化,因此除非事先提供了正确的模式,否则SSPCA算法并不总是能够正确估计所有分量的稀疏模式。

在以上研究的基础上,本文提出了一种自适应稀疏PCA算法(Adaptive sparse principal component analysis, ASPCA)。该算法通过使用组套索模型,在载荷向量上施加块稀疏约束找到最小化的稀疏载荷矩阵和正交矩阵,得出自适应稀疏PCA公式;对稀疏矩阵的不同列使用不同的调整参数以获得自适应惩罚,使得稀疏矩阵不同列具有不同的收缩量,从而在主载荷向量上保留稀疏模式;采用块坐标下降法对自适应稀疏PCA公式进行两阶段优化,找到稀疏载荷矩阵和正交矩阵,提高特征选择的能力。仿真结果表明,相比SPCA算法和SSPCA算法,ASPCA算法的降维效果显著,且能提取更有价值的特征,有力的提高了分类模型的平均分类准确率。

1 稀疏主成分分析

令 X 表示秩为 $q \leq \min(n, p)$ 的 $n \times p$ 数据矩阵,其中 n 为数据样本数, p 为变量数。对于 $i = 1, \dots, n$, X 的行表示为 x_i^{row} ,假定其以0为中心且 $\text{cov}(x_i^{\text{row}}) = \Sigma$, Σ 为 $p \times p$ 的正定矩阵,可以分解为

$$\Sigma = \sum_{i=1}^q \lambda_i \mathbf{v}_i \mathbf{v}_i^{\text{T}} \quad (1)$$

式中: λ_i 表示 Σ 的第 i 个最大特征值; $\mathbf{v}_i = (v_{i1}, \dots, v_{ip})^{\text{T}}$ 为 Σ 关联的特征向量。PCA通过将原始变量替换为主成分 X_{v_k} , $k = 1, \dots, q$ 的 q 个线性组合,从而将数据的维数从 p 减少到 q ,通过最大化其方差获得主载荷向量,即

$$\begin{cases} \mathbf{v}_k = \arg \max_{\mathbf{v}} \text{var}(X_{\mathbf{v}}) \\ \text{s.t. } \mathbf{v}_k^{\text{T}} \mathbf{v}_k = 1; \mathbf{v}_k^{\text{T}} \mathbf{v}_j = 0, j < k \end{cases} \quad (2)$$

式中: \mathbf{v}_k 为第 k 个主载荷向量; X_{v_k} 的投影为第 k 个主成分; $\text{var}(\cdot)$ 表示随机变量的方差。当使用 $q \ll p$ 的分量来解释数据的大部分差异,进而解释基础数据结构时,PCA最为理想。协方差矩阵 Σ 从数据上进行估计,载荷向量从估计矩阵 $\hat{\Sigma}$ 的特征值-特征向量分解进行估计。

基于投影框架导出PCA的替代公式,其中主载荷矩阵(V 定义为包含主要载荷向量的矩阵)通过最小二乘法估计

$$\begin{cases} \min_V \sum_{i=1}^n \|x_i^{\text{row}} - VV^T x_i^{\text{row}}\|_2^2 \\ \text{s.t.} \quad V^T V = I_q \end{cases} \quad (3)$$

式中 $V = (v_1, \dots, v_q)$ 为 $p \times q$ 主载荷矩阵。

使用 X 的奇异值分解 (Singular value decomposition, SVD) 计算 PCA 为

$$X = UDV^T \quad (4)$$

式中: $Z \triangleq UD$ 的列是主成分;而 V 的列是相应的载荷向量。矩阵 D 是有序奇异值 $d_1 \geq d_2 \geq \dots \geq d_q > 0$ 的 $q \times q$ 对角矩阵,并且 U 和 V 的列正交,因此 $U^T U = V^T V = I_q$ 。式(4)的SVD也提供了与 X 最接近的秩 q 矩阵近似,其中 X 和秩 q 近似矩阵 \hat{X}_q 的接近度用F-范数的平方来表示,定义为

$$\|X - \hat{X}_q\|_F^2 = \text{tr} \left\{ (X - \hat{X}_q)(X - \hat{X}_q)^T \right\} \quad (5)$$

这个秩 q 近似也对应于 q 个秩 1 近似,即

$$\min_{u, b} \|X - ub^T\|_F^2 \quad (6)$$

式中: $b = d_1 v$, 向量 u 和 v 分别为 $n \times 1$ 和 $p \times 1$ 的矩阵,且 $\|u\|_2 = \|v\|_2 = 1$ 。本文用式(6)描述PCA的稀疏性。在 b 上添加 L_1 -范数作为稀疏正则约束获得稀疏载荷向量,表示为

$$\|b\|_1 = \sum_{i=1}^p |b_i| \quad (7)$$

式中 b_i 为 b 的分量。在式(7)中,稀疏惩罚对向量 b 的各个元素施加约束,这种约束 b 元素的方式导致了在不同成分之间稀疏模式变化的问题。但现有的稀疏PCA方法均不能保证在所有主载荷向量上都保持稀疏模式,也就是非零变量的位置在不同的主载荷向量上是不同的。

2 自适应稀疏主成分分析算法

2.1 自适应稀疏PCA公式

为了在所有主载荷向量上保留稀疏模式,本文采用组套索模型惩罚分组变量^[10-12]。通过在载荷向量上施加块稀疏约束并找到最小化的向量 U 和 B^T 的列向量,得出自适应稀疏PCA方法的公式为

$$\begin{cases} \min_{U, B} \|X - UB^T\|_F^2 + \sum_{i=1}^p \alpha_i \sqrt{q} \|b_i\|_2 \\ \text{s.t.} \quad U^T U = I_q \end{cases} \quad (8)$$

式中: b_i 为与 $q \times p$ 矩阵 B^T 的第 i 列相对应的 $q \times 1$ 向量,其中 \sqrt{q} 用于相对于向量 b_i 的维数来调整惩罚。式(8)的第1项对应于式(5)的F-范数。式(8)中将 B^T 的每一列视为1个组, $\|b_i\|_2 = 0$ 相当于将 B^T 的第 i 列设置为零。因此, $\|\cdot\|_2$ 惩罚使得 B^T 按列稀疏。式(8)利用了 $b_i = 0$ 时 $\sum_{i=1}^p \|b_i\|_2$ 的不可微性,将所有位于相同位置 i 的 q 系数组设置为零,并在所有载荷向量上实现一致性稀疏。如果保留1个主成分,则列 b_i 为一维,代价函数式(8)为

$$\begin{cases} \min_{u, b} \|X - ub^T\|_F^2 + \sum_{i=1}^p \alpha_i |b_i| \\ \text{s.t.} \quad \|u\|_2 = 1 \end{cases} \quad (9)$$

传统的稀疏PCA是式(8)描述的通用特例,但是如果需要一个以上的主成分来充分描述数据,则本文提出的自适应稀疏PCA方法具有保留相关载荷向量间稀疏模式的优点。式(8)使用 $L_{2,1}$ -范数进行双重正则化,先求矩阵 B^T 第 i 列向量 b_i 的 L_2 范数,再将结果求 L_1 范数,同时实现收缩和选择,可以使一些大小为 q 的系数块恰好为0。代价函数式(8)通过对 B^T 的不同列使用不同的调整参数 α_i 获得自适应惩罚,从而保证 B^T 的不同列出现不同程度的收缩。对于一组 α_i ,式(8)的优化在某种程度上是唯一的,以下命题讨论了式(8)解的唯一性。

命题1 对于给定的秩 q 矩阵 X ,代价函数式(8)的解对 $\hat{U}\hat{B}^T$ 在 $q \times q$ 正交矩阵 Q 上是唯一的。更准确地说,如果 $\hat{U}\hat{B}^T$ 是式(8)的解,则 $\hat{U}Q$ 和 $Q^T\hat{B}^T$ 也是一个解。

证明:假设 (\hat{U}_1, \hat{B}_1) 是式(8)的一个解。对于 $q \times q$ 正交矩阵 Q ,有 $\hat{U}_1\hat{B}_1^T = (\hat{U}_1Q)(Q^T\hat{B}_1^T)$ 且 $\|Q^T b_{1i}\|_2 = \|b_{1i}\|_2$,其中 b_{1i} 是 B_1^T 的第 i 列。因此, $(\hat{U}_2, \hat{B}_2) = (\hat{U}_1Q, \hat{B}_1Q)$ 也是式(8)的解。

假设 (\hat{U}_1, \hat{B}_1) 和 (\hat{U}_2, \hat{B}_2) 是式(8)的两个解。假定 \hat{U}_1 和 \hat{U}_2 为全列秩 q ,存在一个非奇异矩阵 Q 使得 $\hat{U}_2 = \hat{U}_1Q$ 。 \hat{U}_2 的正交性约束 $\hat{U}_2^T\hat{U}_2 = Q^T\hat{U}_1^T\hat{U}_1Q = Q^TQ$ 表示 Q 是正交矩阵,则 $QQ^T = I_q$;因为 $\hat{U}_2(Q^T\hat{B}_1^T) = \hat{U}_1\hat{B}_1$ 和 $\|Q^T b_{1i}\|_2 = \|b_{1i}\|_2$,所以 (\hat{U}_2, \hat{B}_1Q) 也是式(8)的一个解。最后,用 \hat{U}_2 替换式(8)的 U ,则 \hat{B}_2 和 \hat{B}_1Q 都会使凸函数(具有唯一的全局解)最小化,因此 $\hat{B}_2 = \hat{B}_1Q$ 。

命题1指出代价函数式(8)仅取决于 $\text{Range}(\hat{U})$ (\hat{U} 的列跨越的 \mathbb{R}^n 的线性子空间),对任何基础选择都是不变的,在此基础上提出式(8)解的稀疏性模式的命题。

命题2 从代价函数式(8)获得的具有非零 L_2 -范数的 B^T 列是唯一确定的。

证明:假设 $\hat{U}_1\hat{B}_1^T$ 和 $\hat{U}_2\hat{B}_2^T$ 是式(8)的两个解。重新排列 \hat{B}_1^T 的列顺序,使 $\hat{B}_1^T = (\hat{B}_{1n}^T, 0)$,其中 \hat{B}_{1n}^T 不包含所有零项的列。根据命题1,存在一个正交矩阵 Q ,使得 $\hat{B}_2^T = Q^T(\hat{B}_{1n}^T, 0) = (Q^T\hat{B}_{1n}^T, 0)$ 。由于 Q 是正交的,所以 $Q^T\hat{B}_{1n}^T$ 的所有列都不完全为零。

2.2 调整参数计算

在式(8)中,自适应调整参数为惩罚向量 $b_i (i = 1, \dots, p)$ 提供灵活的惩罚,保证不同的 b_i 向量产生不同的收缩量。直观地说,必须对 L_2 范数为零的向量采用较大的收缩量,而对于 L_2 范数非零的向量应使用相对较小的收缩量。在实际中,对 p 个调整参数的选择进行贪婪搜索需要大量的计算。为了减少计算量,提高效率,设 α_i 为

$$\alpha_i = \frac{\alpha}{\|\tilde{d}_i\|_2^\eta} \tag{10}$$

式中 \tilde{d}_i 是从 X 的第1个 q 右奇异向量乘以相关的第1个 q 奇异值且 $\eta = 1$ 得到的未惩罚解。

2.3 代价函数优化

代价函数式(8)是双凸结构,对于固定 U 或 B 都是凸的,采用块坐标下降法^[13]对其进行两阶段优化。第1阶段固定 U 求解 B 对式(8)进行优化,第2阶段固定 B 求解 U 优化式(8)。优化到最后, B 表示稀疏载荷矩阵, U 将是一个完全正交的矩阵。

第1阶段对应于

$$\min_B \|X - UB^T\|_F^2 + \sum_{i=1}^p \alpha_i \sqrt{q} \|b_i\|_2 \tag{11}$$

式(11)的第1项可以改写为

$$\|X - UB^T\|_F^2 = \sum_{i=1}^p \|x_i - Ub_i\|_2^2 \quad (12)$$

式中 x_i 为 X 的第 i 列, 其为 $n \times 1$ 。

关于 b_i 的次梯度方程为

$$-2U^T(x_i - Ub_i) + \alpha_i\sqrt{q}s_i = 0 \quad i = 1, \dots, N \quad (13)$$

式中: 如果 $\|b_i\|_2 \neq 0$, 则 $s_i = \frac{b_i}{\|b_i\|_2}$; 如果 $\|b_i\|_2 = 0$, 则 s_i 是满足 $\|s_i\|_2 < 1$ 的维度 q 的向量。

如果 $\|b_i\|_2 \neq 0$, 式(13)可以直接求解 b_i , 即

$$-2U^T(x_i - Ub_i) + \alpha_i\sqrt{q}\frac{b_i}{\|b_i\|_2} = 0 \quad (14)$$

的解为

$$b_i = \left(1 + \frac{\alpha_i\sqrt{q}}{2\|b_i\|_2}\right)^{-1} U^T x_i \quad (15)$$

对式(15)两边同时取 L_2 范数进行求解

$$\begin{aligned} \|b_i\|_2 &= \frac{\|U^T x_i\|_2}{1 + \frac{\alpha_i\sqrt{q}}{2\|b_i\|_2}} = \frac{\|b_i\|_2 \|U^T x_i\|_2}{\|b_i\|_2 + \frac{\alpha_i\sqrt{q}}{2}} \Rightarrow \\ 1 &= \frac{\|U^T x_i\|_2}{\|b_i\|_2 + \frac{\alpha_i\sqrt{q}}{2}} \Rightarrow \|b_i\|_2 = \left(\|U^T x_i\|_2 - \frac{\alpha_i\sqrt{q}}{2}\right) \end{aligned} \quad (16)$$

将此解代入式(15), 得

$$b_i = \left(1 - \frac{\alpha_i\sqrt{q}}{2\|U^T x_i\|_2}\right) U^T x_i \quad (17)$$

当 $\|b_i\|_2 = 0$ 时, b_i 的次梯度方程变为

$$-2U^T x_i + \alpha_i\sqrt{q}s_i = 0 \quad (18)$$

解 s_i 为

$$s_i = \frac{2}{\alpha_i\sqrt{q}} U^T x_i \quad (19)$$

因此, 如果 $\|s_i\|_2 < 1$ 成立, 则 $\|b_i\|_2 = 0$ 。结合式(17)和式(19)的结果得出式(9)的解为

$$b_i = \left(1 - \frac{\alpha_i\sqrt{q}}{2\|U^T x_i\|_2}\right)_+ U^T x_i \quad (20)$$

对应于向量软阈值规则, 当 $\frac{2}{\alpha_i\sqrt{q}}\|U^T x_i\|_2 < 1$ 时, 运算符 $(\cdot)_+$ 设置为 0。根据式(20), 通过应用一

组软阈值规则来获得 B^T 的不同列, 从而得到式(8)关于 B 的最小化值。

式(8)中优化的第2阶段定义为

$$\min_U \|X - UB^T\|_F^2 \quad \text{s.t. } U^T U = I_q \quad (21)$$

这是一个正交的 Procrustes 问题^[14]。解为 $U = \tilde{U}\tilde{V}^T$, \tilde{U} 和 \tilde{V} 是从 $XB = \tilde{U}\tilde{V}^T$ 的 SVD 得到的, 其中 \tilde{U} 为 $n \times q$, \tilde{V} 为 $p \times q$ 。

2.4 算法描述

根据命题 1 和命题 2 对自适应稀疏 PCA 公式解的说明以及调整参数的计算公式, 利用块坐标下降法两阶段优化的结果, ASPCA 算法伪代码如下所示。

算法 1 自适应稀疏 PCA 算法 (ASPCA)

输入 数据矩阵 X , 稀疏主成分的数量 q , 调整参数 α , 阈值 ϵ 。

输出 稀疏载荷矩阵 B , 正交矩阵 U 。

(1) B 作为第 1 个 q 右奇异向量乘以 X 的第 1 个 q 奇异值;

(2) 使用式(10)计算调整参数 α_{ij} 的向量;

(3) 使用式(20)更新 B^T 及其每列的稀疏性;

(4) For $j = 1 \sim p$

$$(5) \mathbf{b}_i = \left(1 - \frac{\alpha_i \sqrt{q}}{2 \|U_{j-1}^T \mathbf{x}_i\|_2} \right)_+ U_{j-1}^T \mathbf{x}_i;$$

(6) 使用 $XB = \tilde{U}\tilde{V}^T$ 的 SVD 更新 $U_j = \tilde{U}\tilde{V}^T$;

(7) 重复步骤(3~6), 当 $\|U_j - U_{j-1}\|_F < \epsilon$ 时, 停止迭代, 得到更新后的 B^T 和 U 。

通过算法 1 可知, 将数据矩阵、稀疏主成分的数量、调整参数以及阈值作为输入端, 首先计算出调整参数, 然后迭代计算得出矩阵 B^T 和 U , 最后得到稀疏载荷矩阵 B 。ASPCA 算法使用块坐标下降法几次迭代后达到合理的收敛容限, 因此该算法是稳定并保证收敛的。

2.5 算法复杂度分析

对于 ASPCA 算法, 复杂性主要取决于迭代次数。对于非迭代部分, 计算调整参数所需的 SVD 的计算复杂度为 $O(np^2)$ 。更新 U , SVD 的计算复杂度为 $O(qn^2)$, 矩阵乘积计算复杂度为 $O(npq)$ 。因此, 矩阵 U 的计算复杂度由 SVD 决定。更新 B , $U^T \mathbf{x}_i$ 的计算复杂度为 $O(nq)$, 使用式(20)计算 \mathbf{b}_i 的计算复杂度为 $O(nqp)$ 。

3 仿真分析

3.1 数据集

与其他类似算法的仿真研究一致, 本文将 ASPCA 算法应用到 3 种 MNIST 数据条件的数字分类任务中来评估算法降维的有效性。MNIST 数据库^[15]的 3 个不同背景条件为: (1) 无背景; (2) 背景中有均匀的随机噪声; (3) 背景中随机出现无法识别的图像。本文使用大小为 28 像素 \times 28 像素、 $p = 784$ 、 $n = 1000$ 的灰度矢量化图像。图像被矢量化和叠加, 形成 $n \times p$ 数据矩阵 X , 其中 p 为每个图像中变量(即像素)的数量。本文主要研究在主成分上保留稀疏模式的有效性, 降维后执行标准的最近邻分类, 只关注 PCA 降维的效果。使用 $n = 1000$ 个训练样本计算主载荷矩阵, 降维后这些样本用于训练分类器。

3.2 准确率分析

(1) 无背景时的准确率分析

如图 1 所示, 在无背景情况下, 随着主成分数量的增加, SPCA 算法、SSPCA 算法和 ASPCA 算法的平均分类准确率均逐渐提高且相差不大。图中所示的准确率是测试数据的 10 个不相交子集的平均值。

当主成分数量小于50个时,ASPCA算法和SPCA算法的平均分类准确率一致且高于SSPCA算法,因为没有关于稀疏模式结构的先验信息,SSPCA算法需要花时间确定正确模式。也就是当主成分数量小于50个时,ASPCA算法和SPCA算法的降维性能优于SSPCA算法。当主成分数量大于50个后,3种算法的平均分类准确率一致,达到了88%,即在无背景情况下,随着主成分数量的增加,SPCA算法、SSPCA算法和ASPCA算法降维效果逐渐接近且相差不大。

(2) 背景中有均匀随机噪声的准确率分析

如图2所示,当背景中有均匀的随机噪声时,随着主成分数量的增加,ASPCA算法的平均分类准确率优于SPCA算法和SSPCA算法且稳定。在主载荷上引入稀疏度可以帮助选择原始空间中的重要特征,但当提取的主载荷数量增多时,SPCA算法需要分别处理主成分。因此随着主成分数量的增加,SSPCA算法和ASPCA算法的降维效果优于SPCA算法。当主成分数量大于等于90时,SSPCA算法和ASPCA算法的平均分类准确率逐渐相近,达到75.3%,降维性能接近一致。

(3) 背景中随机出现无法识别图像的准确率分析

如图3所示,在背景中随机出现无法识别图像的情况下,随着主成分数量的增多,ASPCA算法、SPCA算法和SSPCA算法的平均分类准确率逐渐提高。因为ASPCA算法能够从数字数据的固有可变性中提取更多信息,从而不会受到背景噪声的干扰,能用更少的主成分实现更高的分类精度。对比SPCA算法和SSPCA算法,ASPCA算法的平均分类准确率较高,即在背景中随机出现无法识别图像的情况下,ASPCA算法的降维效果优于SPCA算法和SSPCA算法。

可见,在对3种MNIST数据条件的数字进行分类的过程中,ASPCA算法的降维效果优于SPCA算法和SSPCA算法,尤其是在存在噪声情况下提取具有相同稀疏性模式的载荷向量是重要的。

4 结束语

为了解决主载荷稠密的问题,本文提出了一种ASPCA算法。通过在载荷向量上施加块稀疏约束

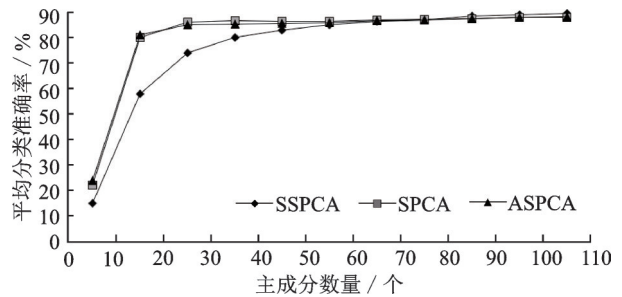


图1 无背景时准确率随主成分数量的变化

Fig.1 Variation of average classification accuracy with the number of principal components in no background

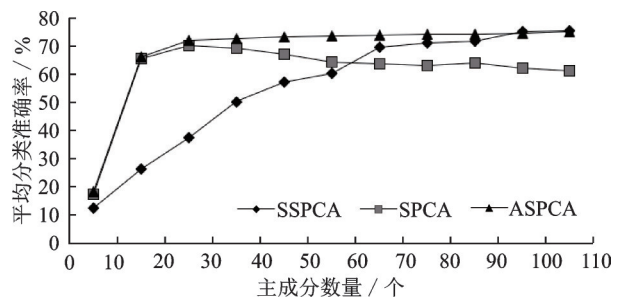


图2 背景中有随机噪声时准确率随主成分数量的变化

Fig.2 Variation of average classification accuracy with the number of principal components in the background with random noise

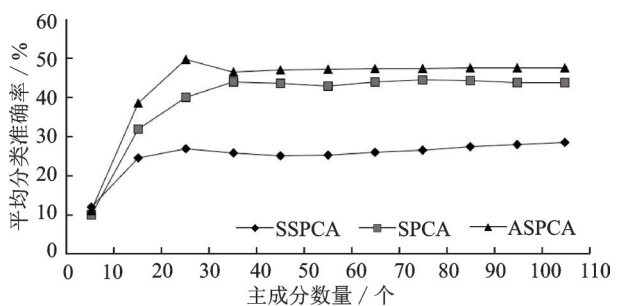


图3 背景中随机出现无法识别图像时准确率随主成分数量的变化

Fig.3 Variation of average classification accuracy with the number of principal components in the background appearing unrecognizable images randomly

得出 ASPCA 公式,对稀疏矩阵的不同列使用不同的调整参数获得自适应惩罚,使主载荷向量在所有主成分上都具有相同的稀疏性模式;然后,通过块坐标下降法进行优化找到最小化的稀疏载荷矩阵和正交矩阵,实现降维的最优化。仿真结果表明,ASPCA 算法的降维性能更优,能提取更有价值的特征,从而有力地提高了分类模型的准确率。

参考文献:

- [1] 谢昆明, 罗幼喜. 一种改进的主成分分析特征抽取算法: YJ-MICPCA[J]. 武汉科技大学学报, 2019, 42(3): 220-226.
XIE Kunming, LUO Youxi. An improved PCA algorithm for feature extraction: YJ-MICPCA[J]. Journal of Wuhan University of Science & Technology, 2019, 42(3): 220-226.
- [2] JOLLIFFE I T, CADIMA J. Principal component analysis: A review and recent developments[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2016, 374(2065): 1-16.
- [3] VASWANI N, CHELLAPPA R. Principal components null space analysis for image and video classification[J]. IEEE Transactions on Image Processing, 2006, 15(7): 1816-1830.
- [4] PYATYKH S, HESSER J, ZHENG L. Image noise level estimation by principal component analysis[J]. IEEE Transactions on Image Processing, 2013, 22(2): 687-699.
- [5] 刘作军, 高尚兵. 基于结构化低秩表示和低秩投影的人脸识别算法[J]. 计算机工程与科学, 2018, 40(1): 108-115.
LIU Zuojun, GAO Shangbing. Face recognition based on structured low rank representation and low rank projection[J]. Computer Engineering & Science, 2018, 40(1): 108-115.
- [6] MIN Feng, YE Xianyi, ZHANG Yanduo. Method of hand written digit recognition based on improved PCANet[J]. Huazhong University of Science & Technology (Natural Science Edition), 2018, 46(12): 101-105.
- [7] HASTIE T, TIBSHIRANI R, WAINWRIGHT M. Statistical learning with sparsity: The LASSO and generalizations[M]. London: Chapman & Hall, 2015.
- [8] ZOU H, HASTIE T, TIBSHIRANI R. Sparse principal component analysis[J]. Journal of Computational and Graphical Statistics, 2006, 15(2): 265-286.
- [9] JENATTON R, AUDIBERT J Y, BACH F. Structured variable selection with sparsity-inducing norms[J]. Journal of Machine Learning Research, 2011, 12: 2777-2824.
- [10] RAO N, NOWAK R, COX C, et al. Classification with the sparse group LASSO[J]. IEEE Transactions on Signal Processing, 2016, 64(2): 448-463.
- [11] CATALINA A, ALAÍZ C M, DORRONSORO J R. Accelerated block coordinate descent for sparse group lasso[C]// Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN). Rio de Janeiro, Brazil: IEEE, 2018.
- [12] WANG H, LENG C. A note on adaptive group lasso[J]. Computational Statistics & Data Analysis, 2008, 52(12): 5277-5286.
- [13] 薛芳. 组 Lasso 模型及坐标下降算法研究[D]. 秦皇岛: 燕山大学, 2015.
XUE Fang. Study on group Lasso model and coordinate descent algorithm[D]. Qinhuangdao: Yanshan University, 2015.
- [14] GOLUB G H, VAN LOAN C F. Matrix computations[M]. [S.l.]: Johns Hopkins University Press, 2012.
- [15] LAROCHELLE H, ERHAN D, COURVILLE A, et al. An empirical evaluation of deep architectures on problems with many factors of variation[C]// Proceedings of the 24th International Conference on Machine Learning. [S.l.]: ACM, 2007: 473-480.

作者简介:



邵剑飞(1970-),男,副教授,硕士生导师,研究方向:图像处理、数据挖掘, E-mail:siee@163.com。



浦蓉(1994-),通信作者,女,硕士,研究方向:图像处理、数据分析, E-mail:1661210225@qq.com。



黄伟(1995-),男,助理工程师,研究方向:数据挖掘。



季建杰(1998-),男,硕士研究生,研究方向:图像处理、数据分析。



郭鹏(1996-),男,硕士研究生,研究方向:图像处理、数据分析。