

Accepted Manuscript

On the orthogonal distance to class subspaces for high-dimensional data classification

Rui Zhu, Jing-Hao Xue

PII: S0020-0255(17)30835-6
DOI: [10.1016/j.ins.2017.07.019](https://doi.org/10.1016/j.ins.2017.07.019)
Reference: INS 12985



To appear in: *Information Sciences*

Received date: 21 February 2017
Revised date: 9 June 2017
Accepted date: 12 July 2017

Please cite this article as: Rui Zhu, Jing-Hao Xue, On the orthogonal distance to class subspaces for high-dimensional data classification, *Information Sciences* (2017), doi: [10.1016/j.ins.2017.07.019](https://doi.org/10.1016/j.ins.2017.07.019)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

On the orthogonal distance to class subspaces for high-dimensional data classification

Rui Zhu^a, Jing-Hao Xue^{a,*}

^a*Department of Statistical Science, University College London, London WC1E 6BT, UK*

Abstract

The orthogonal distance from an instance to the subspace of a class is a key metric for pattern classification by the class subspace-based methods. There is a close relationship between the orthogonal distance and the residual standard deviation of a test instance from the class subspace. In this paper, we shall show that an established and widely-used relationship, between the residual standard deviation and the sum of squares of the residual PC scores, is not precise, and thus can lead to incorrect results, for the inference of high-dimensional data which nowadays are common in practice.

Keywords: Classification, high-dimensional data, orthogonal distance, principal component analysis (PCA), soft independent modelling of class analogy (SIMCA).

1. Introduction

2 In class subspace-based classification methods, a subspace is first learned
3 in the training phase for each class separately from its training data. Then in

*Corresponding author.

Email addresses: r.zhu.12@alumni.ucl.ac.uk (Rui Zhu), jinghao.xue@ucl.ac.uk (Jing-Hao Xue)

Download English Version:

<https://daneshyari.com/en/article/4944310>

Download Persian Version:

<https://daneshyari.com/article/4944310>

[Daneshyari.com](https://daneshyari.com)