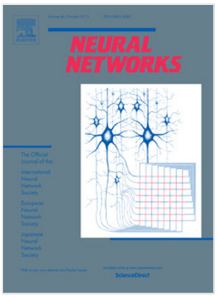
Accepted Manuscript



An online incremental orthogonal component analysis method for dimensionality reduction

Tao Zhu, Ye Xu, Furao Shen, Jinxi Zhao

PII: S0893-6080(16)30140-X
DOI: http://dx.doi.org/10.1016/j.neunet.2016.10.001
Reference: NN 3672

To appear in: *Neural Networks*

Received date: 29 May 2016
Revised date: 17 August 2016
Accepted date: 4 October 2016

Please cite this article as: Zhu, T., Xu, Y., Shen, F., & Zhao, J. An online incremental orthogonal component analysis method for dimensionality reduction. *Neural Networks* (2016), http://dx.doi.org/10.1016/j.neunet.2016.10.001

An Online Incremental Orthogonal Component Analysis Method for Dimensionality Reduction

Tao Zhu[a], Ye Xu[b,1], Furao Shen[a,*], Jinxi Zhao[a]

[a] *National Key Laboratory for Novel Software Technology, Department of Computer Science and Technology, Nanjing University, Nanjing, P.R. China*
[b] *6211 Sudikoff Lab, Dartmouth College Hanover, NH 03755, U.S.*

Abstract

In this paper, we introduce a fast linear dimensionality reduction method named incremental orthogonal component analysis (IOCA). IOCA is designed to automatically extract desired orthogonal components (OCs) in an online environment. The OCs and the low-dimensional representations of original data are obtained with only one pass through the entire dataset. Without solving matrix eigenproblem or matrix inversion problem, IOCA learns incrementally from continuous data stream with low computational cost. By proposing an adaptive threshold policy, IOCA is able to automatically determine the dimension of feature subspace. Meanwhile, the quality of the learned OCs is guaranteed. The analysis and experiments demonstrate that IOCA is simple, but efficient and effective.

Keywords: dimensionality reduction, orthogonal component, incremental learning, automatic target dimension estimation, online learning.

1. Introduction

How to efficiently and effectively extract useful information from high-dimensional data is an open problem worth studying. One of the biggest challenges is "the curse of dimensionality". It is caused by the high dimensional form of original

*Corresponding author
Email addresses: tao144@gmail.com (Tao Zhu), ye@cs.dartmouth.edu (Ye Xu), frshen@nju.edu.cn (Furao Shen), jxzhao@nju.edu.cn (Jinxi Zhao)
[1] This work was done when the author was in Nanjing University.