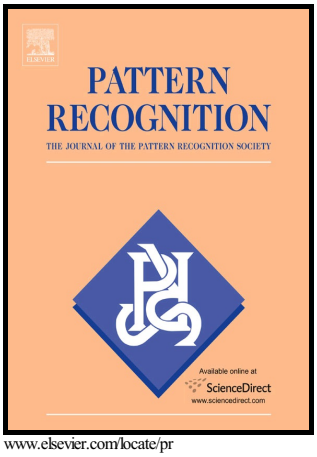
Multi-Manifold Matrix Decomposition for Data Co-clustering

Kais Allab, Lazhar Labiod, Mohamed Nadif



www.elsevier.com/locate/pr

PII: S0031-3203(16)30385-5
DOI: http://dx.doi.org/10.1016/j.patcog.2016.11.027
Reference: PR5970

To appear in: *Pattern Recognition*

Received date: 24 March 2015
Revised date: 25 April 2016
Accepted date: 29 November 2016

Cite this article as: Kais Allab, Lazhar Labiod and Mohamed Nadif, Multi Manifold Matrix Decomposition for Data Co-clustering, *Pattern Recognitio* http://dx.doi.org/10.1016/j.patcog.2016.11.027

Multi-Manifold Matrix Decomposition for Data Co-clustering

Kais Allab [1], Lazhar Labiod, Mohamed Nadif

(1): kais.allab@parisdescartes.fr
LIPADE, University of Paris Descartes. 45, rue des Saints-Peres. Paris, France.
Tel: +33-644-079-623; Fax: +33-144 553-536

Abstract

We propose a novel Multi-Manifold Matrix Decomposition for Co-clustering (M3DC) algorithm that considers the geometric structures of both the sample manifold and the feature manifold simultaneously. Specifically, multiple candidate manifolds are constructed separately to take local invariance into account. Then, we employ multi-manifold learning to approximate the optimal intrinsic manifold, which better reflects the local geometrical structure, by linearly combining these candidate manifolds. In M3DC, the candidate manifolds are obtained using various manifold-based dimensionality reduction methods. These methods are based on different rationales and use different metrics for data distances. Experimental results on several real data sets demonstrate the effectiveness of our proposed M3DC.

Keywords: Co-clustering, Matrix Tri-Factorization, Muli-Manifold learning.

1. Introduction

Clustering is a fundamental topic in several areas including computer vision. Non-negative Matrix Factorization (NMF) [1] has become one of the most frequently used in this context. NMF was proposed to learn a parts-based representation, but it focuses on unilateral clustering i.e. on only one of the two sets of samples or features of a data matrix. There are many different Co-clustering approaches fulfilling this task; see for instance [2, 3] [4, 5]. In this paper, we focus on the Non-negative Matrix Tri-Factorization (NMTF) [6, 7].