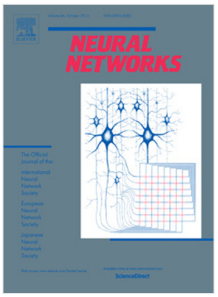
Accepted Manuscript



Graph-based composite local Bregman divergences on discrete sample spaces

Takafumi Kanamori, Takashi Takenouchi

PII: S0893-6080(17)30141-7
DOI: http://dx.doi.org/10.1016/j.neunet.2017.06.005
Reference: NN 3770

To appear in: *Neural Networks*

Received date : 6 January 2017
Revised date : 1 May 2017
Accepted date : 8 June 2017

Please cite this article as: Kanamori, T., Takenouchi, T., Graph-based composite local Bregman divergences on discrete sample spaces. *Neural Networks* (2017), http://dx.doi.org/10.1016/j.neunet.2017.06.005

Graph-based Composite Local Bregman Divergences on Discrete Sample Spaces

Takafumi Kanamori[a], Takashi Takenouchi[b]

[a]*Nagoya University*
RIKEN Center for Advanced Intelligence Project
Furocho, Chikusaku, Nagoya 464-8603, Japan
[b]*Future University Hakodate*
RIKEN Center for Advanced Intelligence Project
116-2 Kamedanakano, Hakodate, Hokkaido, 040-8655, Japan

Abstract

This paper develops a general framework of statistical inference on discrete sample spaces, on which a neighborhood system is defined by an undirected graph. The scoring rule is a measure of the goodness of fit for the model to observed samples, and we employ its localized version, local scoring rules, which does not require the normalization constant. We show that the local scoring rule is closely related to a discrepancy measure called composite local Bregman divergence. Then, we investigate the statistical consistency of local scoring rules in terms of the graphical structure of the sample space. Moreover, we propose a robust and computationally efficient estimator based on our framework. In numerical experiments, we investigate the relation between the neighborhood system and estimation accuracy. Also, we numerically evaluate the robustness of localized estimators.

Keywords: scoring rule, Bregman divergence, coincidence axiom, locality, robustness

1. Introduction

This paper proposes a general framework of statistical inference using unnormalized statistical models on discrete sample spaces. One of the most common methods in statistical inference is the maximum likelihood estimator (MLE), which is obtained by maximizing the empirical mean for the log-likelihood of the statistical model. The MLE has some nice properties such as the statistical consistency and efficiency. The computation of the normalization constant in the statistical model is, however, often intractable in high-dimensional sample domains.

Email addresses: kanamori@is.nagoya-u.ac.jp (Takafumi Kanamori), ttakashi@fun.ac.jp (Takashi Takenouchi)