Accepted Manuscript

General Gaussian estimation

Tonglin Zhang

PII: S0047-259X(18)30111-8
DOI: https://doi.org/10.1016/j.jmva.2018.09.010
Reference: YJMVA 4420

To appear in: *Journal of Multivariate Analysis*

Received date : 21 February 2018

Please cite this article as: T. Zhang, General Gaussian estimation, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.09.010

General Gaussian estimation

Tonglin Zhang*

Department of Statistics, Purdue University, 250 North University Street, West Lafayette, IN 47907, USA

Abstract

This article proposes a general Gaussian estimation approach for situations where the implementation of maximum likelihood estimation is difficult. The primary task is to construct a Gaussian estimation function by putting exact expressions of the mean vector and the variance-covariance matrix of the response vector into the log-likelihood function of the multivariate normal distribution. A Gaussian estimator is derived by maximizing the Gaussian estimation function. This construction can induce an optimality condition that the true parameter vector is the unique maximizer of the expected value of the Gaussian estimation function. The optimality condition is equally important to that given by the Kullback–Leibler information number in the maximum likelihood approach. It is a major condition in the derivation of nice theoretical properties such as consistency and asymptotic normality. The general Gaussian estimation approach can significantly reduce the computational burden when the log-likelihood function of a statistical model contains intractable high-dimensional integrals. By applying it to the Poisson-lognormal model, a special case of generalized linear mixed effect models, a closed-form (i.e., without intractable integrals) estimation approach for fixed effect parameters and variance components is derived. The simulation study shows that the resulting estimator is precise, reliable, and computationally efficient.

Keywords: closed-form estimation, Consistency, General Gaussian estimation, Kullback–Leibler information number, Optimality condition, Poisson-lognormal model
AMS subject classifications: 62F10, 62F12

1. Introduction

The purpose of this article is to propose an approach which can efficiently compute estimates of model parameters when the implementation of maximum likelihood estimation is difficult. It provides new insights to formulate estimation functions with nice theoretical and computational properties when the log-likelihood function of a statistical model contains intractable high-dimensional integrals.

A well-known example is hierarchical models for count data. Suppose a count response is modeled by a few explanatory variables via a generalized linear model (GLM). A generalized linear mixed effects model (GLMM) is derived if random effects appear in its linear component [5]. A GLMM for count data has at least two hierarchical levels. The first level specifies the conditional distribution of the count response given the random effects. The second level specifies the distribution of the random effects. Since dependence is often involved, a multivariate normal distribution is typically used to model the distribution of the random effects. As the likelihood function of the model contains intractable high-dimensional integrals, numerical computation of the maximum likelihood estimator (MLE) of the parameter vector is difficult. The proposed Gaussian estimation approach can partially solve this difficulty. We provide a closed-form (i.e., without intractable integrals) estimation approach to the Poisson-lognormal model, which is an important case of GLMMs for count data. We also extend our findings to a few other models.

The idea of the approach is motivated from the role of the Kullback–Leibler information number in the maximum likelihood approach. Let $\mathbf{y} = (y_1, \ldots, y_n)^\top$ be a response vector with probability density function (PDF) or probability

*Corresponding author
Email address: tlzhang@purdue.edu (Tonglin Zhang)