Journal of Multivariate Analysis 113 (2013) 37–47

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

The multilinear normal distribution: Introduction and some basic properties

Martin Ohlson [a], M. Rauf Ahmad [a], Dietrich von Rosen [a,b,*]

[a] *Department of Mathematics, Linköping University, Sweden*

[b] *Department of Energy and Technology, Swedish University of Agricultural Sciences, Sweden*

ARTICLE INFO

Article history:
Available online 2 June 2011

AMS subject classification:
62H12
62H99

Keywords:
Flip-flop algorithm
Matrix normal distribution
Marginal and conditional distributions
Maximum likelihood estimators
Moments
Tensor product

ABSTRACT

In this paper, the multilinear normal distribution is introduced as an extension of the matrix-variate normal distribution. Basic properties such as marginal and conditional distributions, moments, and the characteristic function, are also presented. A trilinear example is used to explain the general contents at a simpler level. The estimation of parameters using a flip-flop algorithm is also briefly discussed.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The matrix normal distribution, being an extension of the ordinary multivariate (vector-) normal distribution, can be regarded as a *bilinear* normal distribution — a distribution of a two-way (two-component) array, each component representing a vector of observations. The complexity of data, which has become a norm of the day for a variety of applied research areas, requires a consideration of extension of the bilinear normal distribution. The present paper presents this extension, correspondingly named *multilinear* normal distribution [20, Ch. 2], based on a parallel extension of bilinear matrices to multilinear tensors [9]. The adjective *multilinear* has not yet found its way into the general statistical literature. One may, however, trace the same or similar nomenclature with reference to the analysis of complicated data structures, with a commonly used alternative expression being *analysis of multiway data* [21, p. 16]. [21] also gives some useful references on multiway analysis, particularly based on tensor algebra; see also [10].

Compared to the multivariate normal distribution, the multilinear distribution has been a relatively uncharted territory of research. Still, however, some interesting and very useful applications of multilinear distribution can be found in the literature. Particularly, the emergence of complicated and enormous data sets in recent decades has given serious impetus for such applied literature to flourish. As a byproduct, this has caused a huge amount of literature on the theory and applications of tensors in statistics.

One of the most important uses of the multilinear normal (MLN) distribution, and hence tensor analysis, is perhaps in magnetic resonance imaging (MRI). A nice work, particularly focusing on the need to go from matrix-variate to tensor-based

* Corresponding author at: Department of Mathematics, Linköping University, Sweden.
E-mail addresses: martin.ohlson@liu.se (M. Ohlson), muahm@mai.liu.se (M. Rauf Ahmad), dietrich.von.rosen@et.slu.se, Dietrich.von.Rosen@bt.slu.se (D. von Rosen).

0047-259X/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2011.05.015

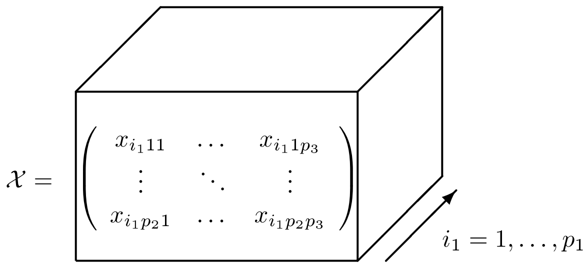

Fig. 1. The box visualizes a three-dimensional data set as a third order tensor.

MLN distribution, is given in [3]. They genuinely argue why a vectorial treatment of a complex data set which actually needs a tensorial treatment and the application of multilinear normality, can lead to wrong or inefficient conclusions. For some more relevant work in the same direction, see [2,4,5], and the references cited therein, whereas a Bayesian perspective is given in [25]; see also [15]. Analysis of multilinear, particularly trilinear data, has a specific attraction in chemometrics and spectroscopy; see for example [23,6]. Other areas of applications include signal processing [18], morphometry [22], geostatistics [24], and statistical mechanics [34], to mention a few. The extensive use of tensor variate analysis in these and other similar fields has generated a special tensorial nomenclature, for example diffusion tensor, dyadic tensor, stress and strain tensors etc. [27]. Similarly, special tensorial decompositions, for example PARAFAC and Tucker decompositions [21], have been developed; for a general comprehensive review of tensor decompositions and their various applications, see [8,19,32].

The use of a tensor, and its associated distributional structure, is even older, and with most frequent applications in the theory of linear models. Some classical treatises on tensors and multilinear algebra are [1,28,7]. For a comprehensive exposition of the use of tensors in statistics, see [27]. In another unique contribution, McCullagh had already introduced tensor notation in statistics with particular reference to the computation of polynomial cumulants [26]; see also [17,11]. The decomposition of ANOVA models into the potential sources of variation is always an important task in the theory of linear models. A tensorial treatment of ANOVA decomposition is given in [36], whereas a study of multilinear skewness and kurtosis in linear models is given in [30]; see also [12]. [14] gives an interesting application in the theory of design of experiments, with particular emphasis on rock magnetism. This paper uncovers some very attractive features of theoretical and geometrical aspects of tensors, when considered from a statistical perspective. The geometrical consideration of tensors in statistics, sometimes even more important than pure theoretical treatment, owes basically to setting the multivariate normality on the Riemannian geometry [33]. As the simplest case of geometrical structure of the parameter space of bivariate normal distribution, see [31], which also uses tensor notation to simplify complicated expressions.

This paper formally introduces MLN distribution, i.e., a normal distribution for the analysis of multiway data, and discusses some basic properties. The rest of the paper is organized as follows. Section 2 introduces the MLN distribution, along with some notation which simplifies the calculations that follow. In Section 3, some properties of the MLN distribution, such as marginal and conditional distributions, moments, and characteristic function, are given. A special case of trilinear normal distribution is interspersed throughout Sections 2 and 3 to explain the theory and notations at a more comprehensible level. Section 4 presents an estimation procedure for the parameters of the distribution.

2. Model

Let $\mathcal{X} = (x_{i_1,\ldots,i_k}) : \times_{i=1}^{k} p_i$ be a tensor of order k, with the dimensions $p_1, p_2, \ldots, p_k$. Fig. 1 shows the special case when $k = 3$. If $p_i = 1$, $2 \le i \le k$ or $3 \le i \le k$ we have the special cases when the tensor equals a vector or a linear mapping.

In order to perform explicit computations, the tensor has to be represented via coordinates. In this paper, the representation will mainly be in vector form. However, the representation of the tensor $\mathcal{X} : \times_{i=1}^{k} p_i$ as a vector can be done in several ways. If we look at the tensor space in Fig. 1, this means that we can look upon the tensor from different directions.

Put $\boldsymbol{e}_{i_1:i_k}^{\boldsymbol{p}} = \boldsymbol{e}_{i_1}^{p_1} \otimes \cdots \otimes \boldsymbol{e}_{i_k}^{p_k}$, where $\boldsymbol{p} = (p_1, \ldots, p_k)$ and $\otimes$ denotes the Kronecker product. To emphasize the dimension, we will write $\boldsymbol{p}_k$, or $\boldsymbol{p}(1 : k)$, instead of $\boldsymbol{p}$. The vectors $\boldsymbol{e}_j^p : p \times 1$ are the unit basis vectors, i.e., a p-vector with 1 in the jth position, and 0 elsewhere. Further, let

$$p_{j:l}^{*} = \prod_{i=j}^{l} p_i \quad \text{and} \quad p_{j:l}^{+} = \sum_{i=j}^{l} p_i, \tag{1}$$

with the special cases

$$p^{*} = p_{1:k}^{*} \quad \text{and} \quad p^{+} = p_{1:k}^{+}, \tag{2}$$

respectively. When there is no ambiguity, we shall drop the dimension from the basis vectors and write $\boldsymbol{e}_{i_1}^{p_1}$ as $\boldsymbol{e}_{i_1}$, and $\boldsymbol{e}_{i_1:i_k}^{\boldsymbol{p}}$ as $\boldsymbol{e}_{i_1:i_k}$, etc. We begin with a formal definition of tensor space.

Download English Version:

https://daneshyari.com/en/article/1145954

Download Persian Version:

https://daneshyari.com/article/1145954

Daneshyari.com