Bayesian Analysis (2011) **6**, Number 2, pp. 197–202

Comment on Article by Hoff

Genevera I. Allen*

1 Introduction

This paper introduces and develops the array normal distribution by extending the matrix-variate normal to the tensor array setting using the Tucker product. Methods for maximum likelihood and Bayesian estimation of separable covariances are given. These contributions are noteworthy as statisticians are encountering increasing numbers of multi-dimensional data sets and methods are needed to model and analyze this array data. Tensor data is especially common in areas of bio-medical imaging, such as neuroimaging and microscopy. With functional magnetic resonance imaging, for example, three-dimensional images of the brain are taken every two to three seconds for many subjects. Often, the dimension of the location variables (voxels) measures in the hundred thousands and the time points measure in the thousands, forming an ultra high-dimensional array. Methods for understanding and modeling these large tensors are certainly needed, and the introduction of the array normal is an important first step in this process.

2 Separable Means

In some cases, having separable means as well as covariances may be useful for tensor data. This may be especially true when no sets of variables along each of the dimensions can be considered independent instances or repeated measures. Then, summing over repeated measures to estimate a mean matrix or mean array is infeasible. Allen and Tibshirani (2010b) modeled separable means for a single instance of a matrix-variate normal data matrix, giving the mean-restricted matrix normal distribution. A similar extension can be employed for the array normal by modeling a separate mean vector for each dimension.

Let $\mathbf{Y}$ be the observed array data, $\mathbf{Y} \in \Re^{m_1 \times \ldots \times m_K}$, and let $\mathbf{M}$ be the mean matrix $\mathbf{M} \in \Re^{m_1 \times \ldots \times m_K}$. Decompose $\mathbf{M} = \sum_{k=1}^{K} \mathbf{M}_k$ where $\mathbf{M}_k = \mathbf{1}_{m_1} \circ \ldots \circ \mathbf{1}_{m_{k-1}} \circ \mu_k \circ \mathbf{1}_{m_{k+1}} \circ \ldots \circ \mathbf{1}_{m_K}$, with $\mu_k \in \Re^{m_k}$, the mean vector of the k^{th} dimension of $\mathbf{Y}$. One can define the mean-restricted array normal as a simple extension of the array normal with the general mean matrix replaced by the structured mean, $\mathbf{M} = \sum_{k=1}^{K} \mathbf{M}_k$. The separable factor means provide a nice analog to the separable covariances of the array normal.

These separable mean parameters can be estimated in a step-wise procedure. (Note that as in the article by Hoff and in other work on the matrix-variate normal (Du-

*Department of Pediatrics-Neurology, Baylor College of Medicine and Department of Statistics, Rice University, Houston, TX, mailto:gallen@rice.edu

 DOI:10.1214/11-BA606A