Reliability Engineering and System Safety 94 (2009) 922–926

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Analyzing the effect of introducing a kurtosis parameter in Gaussian Bayesian networks

P. Main [a,*], H. Navarro [b]

[a] Dpto. Estadística e I.O., Fac. Ciencias Matemáticas, Univ. Complutense de Madrid, 28040 Madrid, Spain
[b] Dpto. de Estadística, I.O. y Cálc. Numérico, Fac. Ciencias, UNED, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 29 October 2007
Received in revised form
5 October 2008
Accepted 21 October 2008
Available online 31 October 2008

Keywords:
Gaussian Bayesian networks
Kullback–Leibler divergence
Exponential power distribution
Sensitivity analysis

ABSTRACT

Gaussian Bayesian networks are graphical models that represent the dependence structure of a multivariate normal random variable with a directed acyclic graph (DAG). In Gaussian Bayesian networks the output is usually the conditional distribution of some unknown variables of interest given a set of evidential nodes whose values are known. The problem of uncertainty about the assumption of normality is very common in applications. Thus a sensitivity analysis of the non-normality effect in our conclusions could be necessary. The aspect of non-normality to be considered is the tail behavior. In this line, the multivariate exponential power distribution is a family depending on a kurtosis parameter that goes from a leptokurtic to a platykurtic distribution with the normal as a mesokurtic distribution. Therefore a more general model can be considered using the multivariate exponential power distribution to describe the joint distribution of a Bayesian network, with a kurtosis parameter reflecting deviations from the normal distribution. The sensitivity of the conclusions to this perturbation is analyzed using the Kullback–Leibler divergence measure that provides an interesting formula to evaluate the effect.

1. Introduction

A Bayesian network (BN) represents a multivariate random variable by means of directed acyclic graphs (DAG) [1,2]. This representation, that describes the dependence structure of the variables, has an intuitive interpretation even for complex models. On the other hand, the quantitative part of the network is given by the corresponding conditional distributions of each variable conditioned by its parents in the graph. Both parts describing the BN can also be used to efficiently make the calculations required for inference about the variables considered. The most important aspect is the evaluation of the effect that evidence about some variables causes to the remainder variables; it is known as evidence propagation. Some algorithms have been implemented giving tools available for inference in BNs [1–4]. These facilities explain the extensive use of BNs in applications [5–7]. In particular they have been found useful as a framework for reliability analysis [8].

In Gaussian BNs (GBN), a joint multivariate normal distribution, $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, is considered. This is usually because the conditionals are also normal distributions and the evidence propagation is obtained directly. However, in practice many problems do not fit this joint model distribution. Thus, in a generalization study, the multivariate exponential power (MEP) distribution [9] could be used. Now, the conditionals are elliptically distributed and the evidence propagation algorithm does not seem to be as direct as the normal case.

Nevertheless, for a given GBN, a sensitivity analysis could provide information about the effect of deviations from normality in the tails. The output to be considered is, as usual in GBN, the conditional distribution of some variables of interest given known values of the evidential variables. Then, we study the difference between the outputs when varying the kurtosis parameter of the MEP family of distributions by means of the Kullback–Leibler (KL) divergence measure [10].

The MEP is a family that extends the corresponding univariate distribution [11]. It has been frequently used for robustness studies because it depends on a kurtosis parameter, β, that goes from zero to infinity, with $\beta = 1$ for the normal distribution, $\beta = \frac{1}{2}$ for the Laplace distribution and $\beta \to \infty$ converging to a uniform distribution. The multivariate extension belongs to the large family of elliptical distributions that includes the multivariate normal. The elliptical continuous distributions are those whose density functions are constant over ellipsoids; this fact explains that many of the nice properties of normal distributions are maintained.

The KL divergence is an asymmetric dissimilarity measure between two probability distributions from an information-theoretic basis. We use it to evaluate the difference between the outputs for a joint distribution with $\beta = 1$ and with some other

* Corresponding author.
E-mail address: pmain@mat.ucm.es (P. Main).

0951-8320/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2008.10.004

$\beta \in (0, \infty)$. The usage of a divergence measure gives a full comparison of the two distributions, not just in some characteristics such as the mean or the variance. With this procedure, if a multivariate normal distribution is used to handle the network but there are some doubts about this model, the sensitivity analysis for the BN reflects the influence of some model deviations on the network output.

Sensitivity analysis in BNs is a developing area of work because its significance in practical situations where the model assumptions are doubtful. Most of the analyses deal with discrete networks [12,13] evaluating the perturbation in the parameters that describe the network. In this respect some sensitivity analyses are studied for GBNs [14] focusing on the effect on the conditional distribution parameters after the evidence propagation. Some other paper [15] uses the KL divergence measure to perform sensitivity analyses in GBNs for perturbations on the parameters of the joint normal distribution of the GBN.

The paper is structured as follows. In Section 2 we set up notation and terminology and review some of the results to be used. In Section 3 our main result is stated and proved. Finally, in Section 4 some relevant examples are analyzed with the proposed methodology and in Section 5 some concluding remarks are given.

2. Preliminaries

Let $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$ be a random variable with a mean vector $\boldsymbol{\mu}$ and a $n \times n$ positive definite matrix $\boldsymbol{\Sigma}$, the MEP$(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \beta)$ is a member of the elliptically contoured family of distributions with a general density function given by

$$h(\mathbf{x}) \propto g((\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}))$$

for the particular case $g^{(\beta)}(t) = \exp\{-t^{\beta}/2\}$. Then the exact density function is

$$f^{(\beta)}(\mathbf{x}) = \frac{n\Gamma\left(\dfrac{n}{2}\right)}{\pi^{n/2}\Gamma\left(1 + \dfrac{n}{2\beta}\right)2^{1+n/2\beta}} |\boldsymbol{\Sigma}|^{-1/2} \times \exp -\frac{1}{2}\{[(\mathbf{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})]^{\beta}\}. \quad (1)$$

Figs. 1–3 show some particular cases for $n = 2$,

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

visualizing the different forms of the densities when the kurtosis parameter β takes some special values which cause a significant variation of the distributions "peakedness".

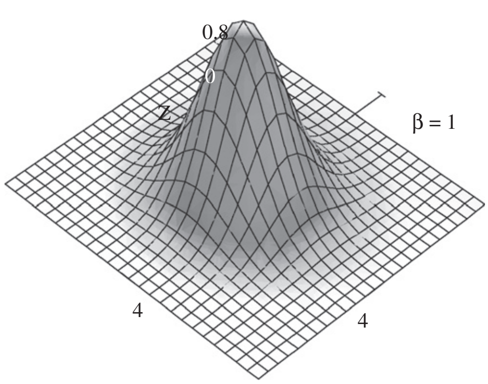

Fig. 1. Multivariate normal density function, $\beta = 1$.

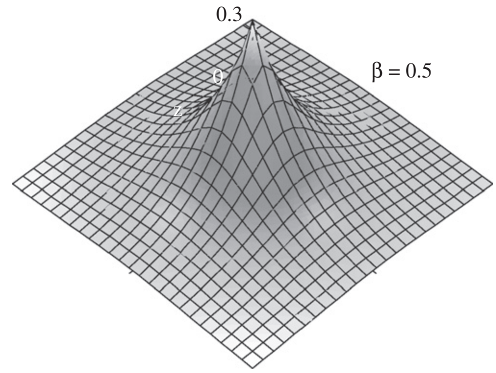

Fig. 2. Multivariate double exponential density, $\beta = \frac{1}{2}$.

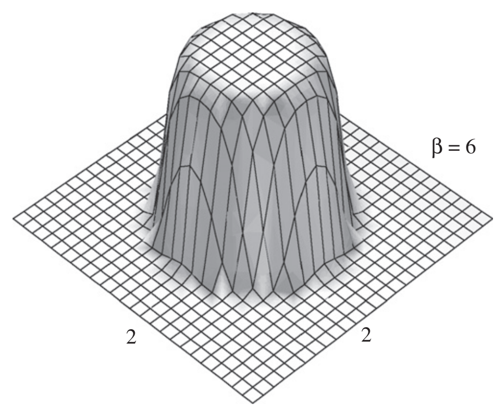

Fig. 3. The MEP density for $\beta = 6$.

It can be shown that for a random variable $\mathbf{X}$ with a density function (1), the mean vector and the covariance matrix are

$$E[\mathbf{X}] = \boldsymbol{\mu}, \quad cov[\mathbf{X}] = \frac{2^{1/\beta}\Gamma\left(\dfrac{n+2}{2\beta}\right)}{n\Gamma\left(\dfrac{n}{2\beta}\right)} \boldsymbol{\Sigma}.$$

The main properties we will use are relative to the conditional distributions but there are many other interesting results for these distributions [9].

2.1. Conditional distributions

Let us consider the partition $\mathbf{X} = (\underbrace{X_1, \ldots, X_p}_{\mathbf{I}}, \underbrace{X_{p+1}, \ldots, X_n}_{\mathbf{E}})^{\mathrm{T}}$ where $\mathbf{I}$ is the subset of variables of interest and $\mathbf{E}$ the evidential variables.

The corresponding partition over the parameters is

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu_i} \\ \boldsymbol{\mu_e} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma_{ii}} & \boldsymbol{\Sigma_{ie}} \\ \boldsymbol{\Sigma_{ei}} & \boldsymbol{\Sigma_{ee}} \end{pmatrix},$$

then, $\mathbf{I}|\mathbf{E} = \mathbf{x}_e$ is elliptically distributed with parameters

$$\boldsymbol{\mu}_{\mathbf{i.e}} = \boldsymbol{\mu}_{\mathbf{i}} + \boldsymbol{\Sigma}_{\mathbf{ie}}\boldsymbol{\Sigma}_{\mathbf{ee}}^{-1}(\mathbf{x_e} - \boldsymbol{\mu}_{\mathbf{e}}),$$
$$\boldsymbol{\Sigma}_{\mathbf{ii.e}} = \boldsymbol{\Sigma}_{\mathbf{ii}} - \boldsymbol{\Sigma}_{\mathbf{ie}}\boldsymbol{\Sigma}_{\mathbf{ee}}^{-1}\boldsymbol{\Sigma}_{\mathbf{ei}},$$
$$g_{\mathbf{i.e}}^{(\beta)}(t) = \exp\{-\tfrac{1}{2}(t + \mathbf{q_e})^{\beta}\},$$

where $\mathbf{q_e} = (\mathbf{x_e} - \boldsymbol{\mu}_{\mathbf{e}})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{ee}}^{-1}(\mathbf{x_e} - \boldsymbol{\mu}_{\mathbf{e}})$ is related to the squared Mahalanobis distance from the evidence to its mean [9]. Then, the functional parameter of the conditional distribution is the