Symbolic Propagation and Sensitivity Analysis in Gaussian Bayesian Networks with Application to Damage Assessment

E. Castillo[1], J. M. Gutiérrez[1], A. S. Hadi[2] and C. Solares[1]

[1] *Department of Applied Mathematics and Computational Sciences, University of Cantabria, SPAIN*

[2] *Department of Statistics, Cornell University, USA*

ABSTRACT

In this paper we show how Bayesian network models can be used to perform a sensitivity analysis using symbolic, as opposed to numeric, computations. An example of damage assessment of concrete structures of buildings is used for illustrative purposes. Initially, normal or Gaussian Bayesian network models are described together with an algorithm for numerical propagation of uncertainty in an incremental form. Next, the algorithm is implemented symbolically, in Mathematica code, and applied to answer some queries related to the damage assessment of concrete structures of buildings. Finally, the conditional means and variances of the of nodes given the evidence are shown to be rational functions of the parameters, thus, discovering its parametric structure, which can be efficiently used in sensitivity analysis.

Key Words: Expert systems, Multivariate normal distribution, Symbolic computations.

1 Introduction

In recent years much attention has been focussed on the use of probabilistic models in expert systems. Today, probabilistic models, especially those associated with Bayesian networks, are gaining more and more popularity as a formalism for handling uncertainty. The increasing number of applications in the last few years also show that this formalism has practical value (an ever growing list of applications in several disciplines: Medicine, Engineering, etc., is available by anonymous FTP from: research.microsoft.com:/pub/dtg/bn-apps.ps).

One of the key problems in Bayesian networks is evidence propagation, which consists of updating the the posterior probabilities of a set of variables of interest whenever a new evidence becomes available. There exist several well known algorithms for the exact and approximate propagation of evidence in Bayesian networks[1−6]. However, from a practical point of view, most of these methods are restrictive because they require all variables to be discrete, while many examples arising in practice involve continuous variables.

On the other hand, propagation algorithms require that the joint probabilities of the nodes be specified numerically, that is, all the parameters must be assigned numeric values. In practice, exact numeric specification of these parameters may not be available or it may happens that the subject matter specialists can specify only ranges of values for the parameters rather than their exact values. In such cases, there is a need for symbolic methods which are able to deal with the parameters themselves, without assigning them values. Symbolic propagation leads to probabilities which are expressed as functions of the parameters instead of real numbers. Thus, the answers to specific queries can then be obtained by plugging the values of the parameters in the solution, without need to redo the propagation. The real practical use of this approach is the possibility of performing a sensitivity analysis of the parameter values without the need of redoing the computations. Symbolic propagation algorithms have been recently introduced to propagate evidence in discrete[7–10] and continuous[11–12] Bayesian networks.

The main contribution of this paper consists of presenting a conceptually simple and efficient algorithm for numeric and symbolic propagation in Gaussian Bayesian networks, discovering the algebraic structure of marginal and conditional probabilities and showing how it can be applied to the damage assessment of reinforced concrete structures of buildings. For the numerical case, we introduce an incremental linear-time algorithm to update probabilities when single pieces of evidence are observed. The capabilities of this method for symbolic computation are also analyzed, showing that the same algorithm can easily be adapted using any standard program with symbolic capabilities. As an illustrative practical example, we use the damage assessment of reinforced concrete structures of buildings.

The paper is organized as follows. In Section 2, we introduce a model to assess the damage of reinforced concrete structures of buildings. Section 3 introduces the Gaussian Bayesian network as a model for continuous random variables. In Sections 4 and 5 a method for both numeric and symbolic propagation is presented and certain questions regarding the assessment of the damage of reinforced concrete structures of buildings are answered. In Section 6 we discuss the algebraic structure of probabilities of single nodes or sets of nodes. Finally, in Section 7 we give some conclusions.

2 Damage Assessment of Buildings

Assessment of the damage of existing buildings is a necessary task to make appropriate strengthening or maintenance plans. Due to the complexity and the uncertainty associated with the lack of knowledge of existing buildings, making such assessment is difficult. In a recent work[13], Liu and Li proposed a model to build an expert system for the assessment of the damage of reinforced concrete structures of buildings. In this paper, we use a slightly modified version of this model, for illustrative purposes.

The model formulation process usually starts with the selection or specification

	Variable	Description
Goal	X_1	Damage assessment
Unobservable	X_2	Cracking state
	X_3	Cracking state in shear domain
	X_4	Steel corrosion
	X_5	Cracking state in flexure domain
	X_6	Shrinkage cracking
	X_7	Worst cracking state in flexure domain
	X_8	Corrosion state
Observable	X_9	Weakness of the beam
	X_{10}	Deflection of the beam
	X_{11}	Position of the worst shear crack
	X_{12}	Breadth of the worst shear crack
	X_{13}	Position of the worst flexure crack
	X_{14}	Breadth of the worst flexure crack
	X_{15}	Length of the worst flexure cracks
	X_{16}	Cover
	X_{17}	Structure age
	X_{18}	Humidity
	X_{19}	PH value in the air
	X_{20}	Content of chlorine in the air
	X_{21}	Shear cracks state
	X_{22}	Flexure cracks state
	X_{23}	Shrinkage
	X_{24}	Corrosion

Table 1: Definitions of the variables related to damage assessment of reinforced concrete structures.

of a set of variables of interest. This specification is dictated by the subject matter specialists. In our example, the goal variable (the damage of a reinforced concrete beam) is denoted by X_1. Another 16 variables $(X_9, X_{10}, \ldots, X_{24})$ have been identified as the main variables influencing the damage of reinforced concrete structures. In addition, the model is built with seven intermediate unobservable conceptual variables $(X_2, X_3, \ldots, X_8)$ which define some partial states of the structure. Table 1 shows the list of variables and their physical meanings. The variables are measured using a scale that is directly related to the goal variable, that is, the higher the value of the variable the more the possibility of damage.

The next step in model formulation is the identification of the dependency structure among the selected variables. In our example, there exist the following cause-effect relationships. The goal variable, X_1, depends primarily on three factors, X_9, the weakness of the beam available in the form of a damage factor, X_{10}, the deflection of the beam, and X_2, its cracking state. The cracking state, X_2, in turn is characterized by four variables: X_3, the cracking state in the shear domain; X_6, the evaluation of the shrinkage cracking; X_4, the evaluation of the steel corrosion; and X_5, the cracking state in the flexure domain. Shrinkage cracking, X_6, depends on shrinkage, X_{23}, and the corrosion state, X_8. Steel corrosion, X_4, is influenced by X_8, X_{24}, and X_5. The cracking state in the shear domain, X_3, depends on X_{11}, the position of the worst shear crack; X_{12}, the breadth of the worst shear crack, X_{21}, the shear cracks state, and X_8. The cracking state in the flexure domain, X_5 is determined by X_{13}, the position of the worst flexure crack, the worst cracking state in the flexure domain without considering the position, X_{22}, the flexure cracks state, and X_7, the worst cracking state in the flexure domain. The variable X_7 is a function of X_{14}, the breadth of the worst flexure crack, X_{15}, the length of the worst flexure crack, X_{16}, the cover, X_{17} the structure age, and X_8, the corrosion state. Node X_8 is determined by X_{18}, the humidity, X_{19}, the PH value in the air, and X_{20}, the content of chlorine in the air.

These cause-effect relationships among the variables are depicted in Figure 1. Each node in this diagram represents a variable. The relationships are represented by links (a directed line emanating from one node and pointing to another). For example, there are three arrows emanating from the nodes X_9, X_{10}, and X_2 and pointing to X_1 indicating that X_1 has three direct causes. The numbers indicated on the links will be explained later.

It is important to notice that the subject matter specialists can develop different dependence structures associated with the same practical problem. Moreover, it is a hard task to develop a consistent and non-redundant probabilistic network. Castillo, Gutiérrez and Hadi[14] have study this problem from a practical viewpoint and describe the steps to be followed in order to generate cause-effect diagrams like the one in Figure 1.

3 Gaussian Bayesian Networks

Let $X = \{X_1, X_2, \ldots, X_n\}$ be a set of n continuous variables and let D be a directed acyclic graph (DAG) with one node for each variable in X (see, for example, Figure 1). The words *node* and *variable* are used synonymously. Every link $X_i \rightarrow X_j$ in the graph indicates a direct dependency between the variables X_i and X_j. The node X_i is called a parent of X_j and X_j is a child of X_i. The set of all parents of a node X_i is denoted as Π_i. For example, in Figure 1, the nodes X_2, X_9, and X_{10} are the parents of X_1, $\Pi_1 = \{X_2, X_9, X_{10}\}$, and X_1 is a child of each of X_2, X_9, and X_{10}. Bayesian network models exploit the topology of a DAG D to define a joint probability density (JPD) consistent with the dependency structure encoded in the graph[15–16].

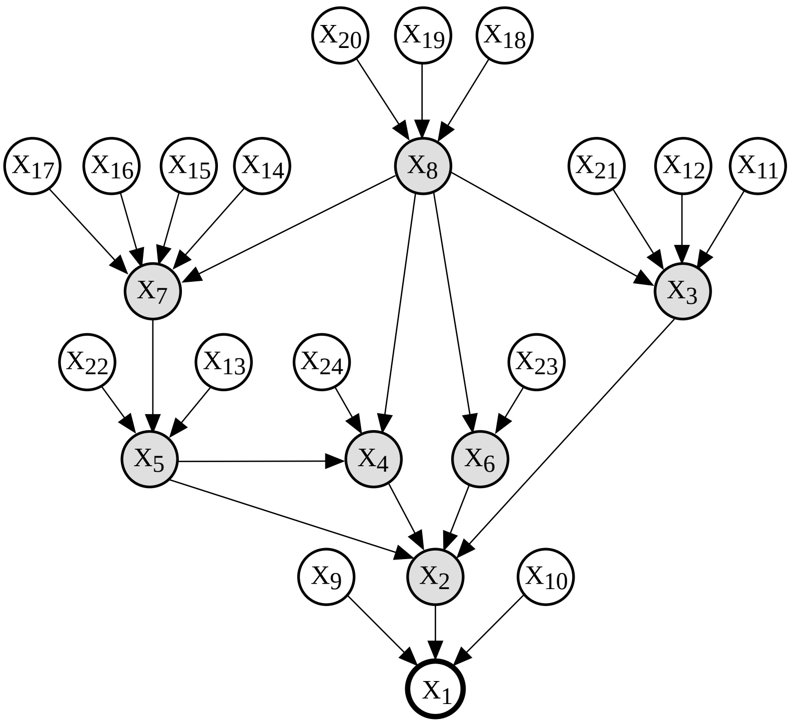

Figure 1: A graph representing the damage assessment of reinforced concrete structure. Shaded nodes represent unobservable (auxiliary) variables.

Assume that the JPD of $\mathbf{X}$ is normal $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, that is,

$$f(\mathbf{x}) = (2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-1/2(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \tag{1}$$

where $\mathbf{x}$ is a realization of the random variable $\mathbf{X}$, $\boldsymbol{\mu}$ is the n-dimensional mean vector, $\boldsymbol{\Sigma}$ is the $n \times n$ covariance matrix, $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$, and $\boldsymbol{\mu}^T$ denotes the transpose of $\boldsymbol{\mu}$. Sometimes it is convenient to refer to the *precision matrix* $\mathbf{W} = \boldsymbol{\Sigma}^{-1}$.

Any multivariate normal JPD function $f(\mathbf{x})$ can be written as a product of conditional probability densities (CPDs) as follows[17]:

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_i(x_i|x_1, \ldots, x_{i-1}), \tag{2}$$

where

$$f_i(x_i|x_1, \ldots, x_{i-1}) \sim N\left(m_i + \sum_{j=1}^{i-1} \beta_{ij}(x_j - m_j),\ \frac{1}{v_i}\right), \tag{3}$$

where m_i is the unconditional mean of X_i, v_i is the conditional variance of X_i, given values for $X_1, \ldots, X_{i-1}$, and β_{ij} is the regression coefficient of X_j when X_i is regressed on $x_1, \ldots, x_{i-1}$.

Gaussian Bayesian networks[18] are introduced as special cases of multivariate normal distributions in which the CPDs in (3) are defined according to a DAG. More formally, a *Gaussian Bayesian network* is a pair (D, P), where

- D is a DAG containing the set of nodes $\{X_1, \ldots, X_n\}$.
- P is a collection of parameters $\mathbf{m} = (m_1, \ldots, m_n)$, $\mathbf{v} = (v_1, \ldots, v_n)$, and $\{\beta_{ij}|j < i\}$, as shown in (3).
- The JPD function of $(X_1, \ldots, X_n)$ is given by (2) where $\beta_{ij} = 0$ if and only if there is no link from X_j to X_i.

Note that in a Gaussian Bayesian network, $\beta_{ij} = 0$ in (3) implies that X_j is not a parent of X_i. This property represents the relationship between the graphical and the probabilistic structure.

Alternatively, we can define the JPD function by giving its mean vector and its covariance matrix. The covariance matrix is symmetric and positive definite. The v_i are positive, and the remaining parameters in (3), β_{ij} and m_j, are arbitrary constants.

Shachter and Kenley [18] describe the general transformation from $\mathbf{v}$ and $\{\beta_{ij}|j < i\}$ to the precision matrix $\mathbf{W}$ of the normal distribution. They use the following recursive formula in which $\mathbf{W}(i)$ denotes the $i \times i$ upper left submatrix of $\mathbf{W}$ and $\boldsymbol{\beta}_i$ denotes the column vector $\{\beta_{ij}|j < i\}$:

$$\mathbf{W}(i+1) = \begin{pmatrix} \mathbf{W}(i) + \dfrac{\boldsymbol{\beta}_{i+1}\boldsymbol{\beta}_{i+1}^T}{v_{i+1}} & \dfrac{-\boldsymbol{\beta}_{i+1}}{v_{i+1}} \\ \dfrac{-\boldsymbol{\beta}_{i+1}^T}{v_{i+1}} & \dfrac{1}{v_{i+1}} \end{pmatrix}, \tag{4}$$

with $\mathbf{W}(1) = 1/v_1$.

Thus, we can consider the DAG in Figure 1 as the network structure of a Gaussian Bayesian network. Then, the next step in Bayesian network formulation is to define a JPD function from (2). For illustrative purposes, we take the value zero for the initial mean of all variables and apply the above method to build the precision matrix $\mathbf{W}$. The coefficients β_{ij} in (3) are shown in Figure 2 and the conditional variances are given by:

$$v_i = \begin{cases} 10^{-4}, & \text{if } X_i \text{ is unobservable.} \\ 1, & \text{otherwise.} \end{cases}$$

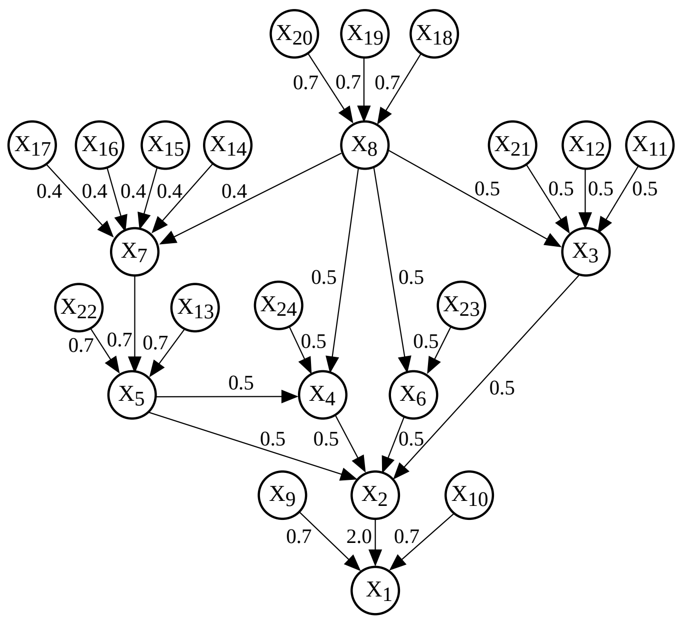

Figure 2: The network in Figure 1 with the coefficient β_{ij} written near the link between X_i and X_j.

4 Numeric Propagation of Uncertainty

In this section we give an algorithm for calculating the updating the probability distributions of the nodes in the network when some evidence is known. The main result is given in the following theorem (see, for example, Anderson[19]).

THEOREM 1 *Suppose that an n-dimensional vector $\mathbf{X}$ has a multivariate normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ with density as in (1). Consider a partition of $\mathbf{X}$ in two sub-vectors $\mathbf{Y} \sim N(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$ and $\mathbf{Z} \sim N(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z)$. Let $\boldsymbol{\Sigma}_{yz}$ be the covariance matrix of $(\mathbf{Y}, \mathbf{Z})$. Then, we have*

$$E[\mathbf{Y}|\mathbf{Z} = \mathbf{z}] = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yz}\boldsymbol{\Sigma}_z^{-1}(\mathbf{z} - \boldsymbol{\mu}_z) \tag{5}$$

and

$$Var[\mathbf{Y}|\mathbf{Z}=\mathbf{z}] = \boldsymbol{\Sigma}_y - \boldsymbol{\Sigma}_{yz}\boldsymbol{\Sigma}_z^{-1}\boldsymbol{\Sigma}_{yz}^T. \tag{6}$$

Let $\mathbf{E} \subset \mathbf{X}$ be a set of evidential variables whose values are known to be $\mathbf{e}$. Theorem 1 suggests an obvious procedure to obtain the mean and variances of any single node given $\mathbf{E} = \mathbf{e}$ (this fact will be used in Section 5 to deal with symbolic propagation). However, it is more convenient to use an incremental method that allows obtaining the conditional distribution of any subset of variables $\mathbf{Y} \subset \mathbf{X}$ as follows. At a given step, replacing $\mathbf{z}$ in (5) and (6) by $\mathbf{e}$, we obtain the mean and covariance matrix of the updated distribution of the non-evidential nodes with a minimum of calculations. It is important to point out that we get the joint distribution of the remaining nodes, which is normal, and then we can answer questions involving the joint distribution of nodes instead of the usual information that refers only to individual nodes.

Note also that if we consider only one evidential node in this step (taking elements one by one from $\mathbf{e}$), we need not calculate the inverse of a matrix because it degenerates to a scalar. In this case $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_{yz}$ are column vectors, and $\boldsymbol{\Sigma}_z$ is a scalar. Then, the number of calculations needed to update the probability distribution of the non-evidential variables, given a single piece of evidence, is linear in the number of variables in $\mathbf{X}$. Thus, this algorithm provides a simple and efficient method for evidence propagation in Gaussian Bayesian networks.

Due to the simplicity of this incremental algorithm, the implementation of this propagation method in the inference engine in an expert system is an easy task. Figure 3 shows the pseudocode to implement this algorithm. The algorithm give the JPD of the non-evidential nodes $\mathbf{Y}$ given the evidence $\mathbf{E} = \mathbf{e}$.

To illustrate the performance of this algorithm we apply it to the damage assessment model introduced in Section 2. We assume that the engineer examines a given concrete beam and obtain the values $x_9, x_{10}, \ldots, x_{24}$ corresponding to the observable variables $X_9, X_{10}, \ldots, X_{24}$. Our aim is to answer certain queries prompted by the engineer about the damage of the beam (the goal variable, X_1). For the sake of simplicity, we consider $x_i = 1,\ i = 9, \ldots, 24$, indicating increasing damage of the beam. In the next example, we show the mean and covariance matrix when applying the incremental algorithm considering the evidences $x_9 = 1, \ldots, x_{24} = 1$. In the last step of the algorithm, that is, when all evidential variables have been considered, the updated normal distributions for the remaining nodes (unobservable and goal nodes, $\mathbf{Y} = (X_8, \ldots, X_1)$), has the following mean and variance matrices:

$$E(\mathbf{Y}|\mathbf{E}=\mathbf{e}) = (2.100, 2.440, 1.550, 3.108, 3.104, 2.550, 5.156, 11.712),$$

$\mathbf{Y} \leftarrow \mathbf{X}$

$\boldsymbol{\mu} \leftarrow E[\mathbf{X}]$

$\boldsymbol{\Sigma} \leftarrow Var[\mathbf{X}]$

For $i \leftarrow 1$ to the number of elements in $\mathbf{E}$, do:

  $\mathbf{z} \leftarrow$ the ith element of $\mathbf{e}$

  $\mathbf{Y} \leftarrow \mathbf{Y} \setminus \mathbf{Z}$

  $\boldsymbol{\mu} \leftarrow \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yz} \boldsymbol{\Sigma}_z^{-1} (\mathbf{z} - \boldsymbol{\mu}_z)$

  $\boldsymbol{\Sigma} \leftarrow \boldsymbol{\Sigma}_y - \boldsymbol{\Sigma}_{yz} \boldsymbol{\Sigma}_z^{-1} \boldsymbol{\Sigma}_{yz}^T$

$f(\mathbf{y}|\mathbf{e} = \mathbf{e}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Figure 3: Incremental algorithm for updating the joint probability of the non-evidential nodes $\mathbf{Y}$ given the evidence $\mathbf{E} = \mathbf{e}$.

$$
Var(Y|\mathbf{E} = \mathbf{e}) = \begin{pmatrix}
X_8 & \ldots & X_4 & X_3 & X_2 & X_1 \\
0.00010 & \ldots & 0.00006 & 0.00005 & 0.00010 & 0.00019 \\
0.00004 & \ldots & 0.00006 & 0.00002 & 0.00009 & 0.00018 \\
0.00005 & \ldots & 0.00003 & 0.00003 & 0.00010 & 0.00020 \\
0.00003 & \ldots & 0.00009 & 0.00001 & 0.00014 & 0.00028 \\
0.00006 & \ldots & 0.00018 & 0.00003 & 0.00017 & 0.00033 \\
0.00005 & \ldots & 0.00003 & 0.00012 & 0.00010 & 0.00020 \\
0.00010 & \ldots & 0.00017 & 0.00010 & 0.00035 & 0.00070 \\
0.00019 & \ldots & 0.00033 & 0.00020 & 0.00070 & 1.00100
\end{pmatrix}.
$$

Note that, in this case, all elements in the covariance matrix but $\Sigma_{1\,1}$ are close to zero indicating that the mean values are quasi-exact estimates for $X_2, \ldots, X_8$ and a good estimation for X_1.

4.1 Answering Queries

The above results can be used to assess the damage (the goal variable X_1) in each of the following hypothetical situations:

Q1: Before Observing Evidence. Initially, we are given the initial mean and covariance matrix introduced in Section 3, without any evidence (i.e., without knowledge of the values $x_9, x_{10}, \ldots, x_{24}$).

A1: Table 2 shows the probabilities of the damage X_1 of a given beam for various types of evidence ranging from no knowledge at all to the knowledge of all the observed values $x_9, x_{10}, \ldots, x_{24}$. Thus, the initial mean and variance of X_1 are 0

Step	Known Variables	Damage of the beam	
		Mean	Variance
0	None	0.000	11.561
1	$X_9 = 1.0$	0.700	11.071
2	$X_{10} = 1.0$	1.400	10.581
3	$X_{11} = 1.0$	1.900	10.331
4	$X_{12} = 1.0$	2.400	10.081
5	$X_{13} = 1.0$	3.450	8.979
6	$X_{14} = 1.0$	3.870	8.802
7	$X_{15} = 1.0$	4.290	8.627
8	$X_{16} = 1.0$	4.710	8.449
9	$X_{17} = 1.0$	5.130	8.273
10	$X_{18} = 1.0$	6.474	6.467
11	$X_{19} = 1.0$	7.818	4.660
12	$X_{20} = 1.0$	9.162	2.854
13	$X_{21} = 1.0$	9.662	2.604
14	$X_{22} = 1.0$	10.712	1.501
15	$X_{23} = 1.0$	11.212	1.251
16	$X_{24} = 1.0$	11.712	1.001

Table 2: Conditional means and variances of the damage, X_1, at the different steps of the incremental algorithm. The *ith* step corresponds the accumulated evidence of the first i evidential variables.

and 11.561, respectively. Other values in Table 2 are explained and interpreted below.

Q2: Observing Some Evidence. Suppose that we observe the value of only one key variable X_9, the weakness of the beam, and it turned out to be $X_9 = 1.0$, an indication that the beam is weak.

A2: Propagating uncertainty with the evidence $X_9 = 1.0$ gives: $E(X_1|X_9 = 1) = 0.70$ and $Var(X_1|X_9 = 1) = 11.071$. Note that after observing the evidence x_9, the mean has increased from 0 to 0.7 and the variance of X_1 has decreased from 11.561 to 11.071.

Q3: Observing Multiple Evidence. Now, suppose that we have the data for all the observable variables as given in Table 2, but the data are measured sequentially.

A3: The answer is given in Table 2, where the probabilities in the *ith* row is computed using the incremental algorithm in the order given in the table, that is, they are based on accumulated evidence. For example, as can be seen in the last row of the table, when all the evidences are considered, $E(X_1|X_9 = 1, \ldots, X_{24} = 1) = 11.712$ and $Var(X_1|X_9 = 1, \ldots, X_{24} = 1) = 1.001$, an indication that the building is seriously damaged. Figure 4 shows several of the updated normal conditional distributions of node X_1, when a new evidence is considered. The figure shows the increasing damage of the beam at different steps, which is indicated by increasing conditional means and decreasing conditional variances.

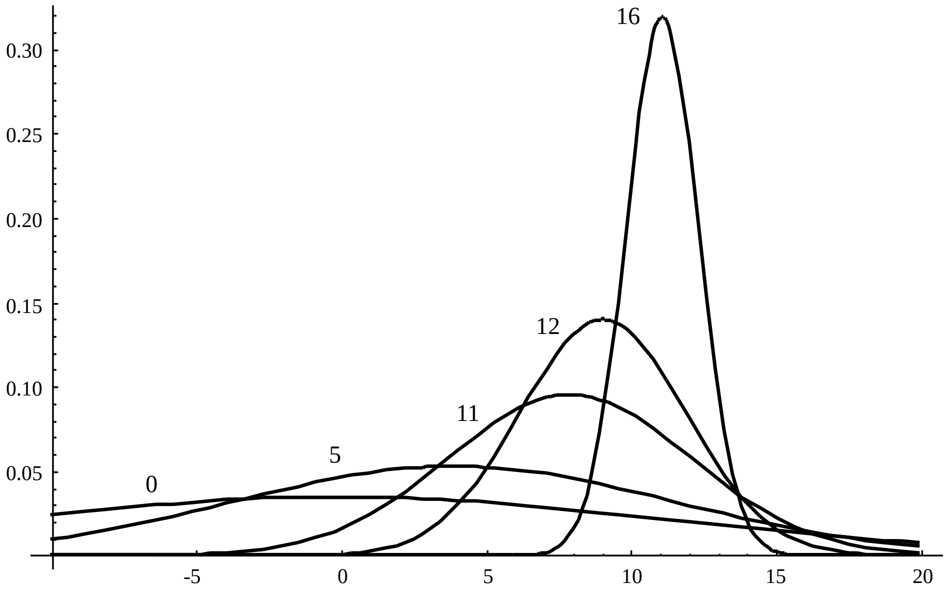

Figure 4: Conditional distributions of node X_1 at various steps of the incremental algorithm. The number on each curve indicate the step number of the incremental algorithm given in Table 2.

It can be seen from the above examples that any query posed by the engineer can be answered simply by propagating evidence using the incremental algorithm. An advantage of the incremental inference algorithm is that we may be able to make a decision concerning the state of damage of a given building immediately after observing only a subset of the variables.

Another important advantage is that, at each step, one can choose the variable for which the evidence should be obtained next in an optimal way; that is, one can use (5) and (6) to choose the variable which will provide the most valuable information about the goal variable. For example, suppose at a given step we can obtain the evidence for one of several variables, which variable should we choose? The answer is the one which is expected to change the conditional mean of the goal variable the most. From equation (5), the change is given by $\boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_z^{-1}(\mathbf{z} - \boldsymbol{\mu}_z)$. Thus, one can calculate the associated changes in the conditional means of the candidate variables then select the variable with the largest change.

5 Symbolic Computations

Dealing with symbolic computations is the same as dealing with numeric values with the only difference being that all the required operations must be performed by a program with symbolic manipulation capabilities. The resulting symbolic expressions provide a useful information both for directly obtaining numerical solutions for different combinations of parameter values and for performing sensitivity analysis. Symbolic computations, however, are intrinsically slow and require more memory.

Consider the damage assessment Bayesian network with the joint probability given in Section 3. Suppose we are interested in the influence that the deflection of the beam, X_{10}, has on its damage assessment, X_1. Then, we can consider X_{10} as a symbolic node modifying the initial mean and covariance matrices. Let $E(X_{10}) = m$, $Var(X_{10}) = v$, and $Cov(X_1, X_{10}) = c$.

We use Mathematica[20] for the symbolic implementation of the propagation algorithm. The resulting code is shown in Figure 5. The program calculates the means and variances of all nodes given the evidence in the evidence list.

Table 3 shows the symbolic results obtained from the propagation of evidence in the damage assessment Bayesian network. This table shows the initial probability of X_1, and the same probability after each one of the evidences, $X_9 = 1, X_{10} = 1, X_{11} = x_{11}, X_{12} = 1, X_{13} = x_{13}, X_{14} = 1$, is sequentially known. Note that some of the evidences have been given in a symbolic form. An examination of the results in Table 3 shows that the conditional means and variances are rational expressions, that is, quotients of polynomials in the parameters. Note that the polynomials are first degree in m, v, x_{11} and x_{13}, that is, in the mean and variance parameters, and in the evidence variables, and second degree in c, the covariance parameter. Note also the common denominator for the rational functions giving the conditional mean and the conditional variance. The fact that the mean and variances of the conditional probability distributions of the nodes are rational functions with polynomials of the indicated degrees is proven in Section 6.

Note that the values in Table 2 are a special case of the one in this example. They can be obtained by setting $m = 0$, $v = 1$ and $c = 0.7$ and considering the evidence values $x_{11} = 1, x_{13} = 1$. Thus the means and variances in Table 2 can actually be obtained from Table 3 by replacing the parameters by their values. For example, for the case of the evidence $X_9 = 1, X_{10} = 1, X_{11} = x_{11}$, the conditional mean of X_1, $(c - cm + 0.7v + 0.5vx_{11})/v$, takes the value 1.9 which is the same result shown in Table 2. Similarly, the conditional variance of X_1 is $(-c^2 + 10.821v)/v = 10.331$ as shown in Table 2.

```
(* ‘Mean’ and ‘Variance’ are given in Section 3 *)

(* Evidence order *)
Evidence=Range[9,24]; EviValue=Table[1,{15}];

(* Conditional Probabilities *)
For[k=0,k<=Length[Evidence],k++,
  For[i=1,i<=Length[Mean],i++,
    If[MemberQ[Take[Evidence,k],i],
      condmean=x[i];
      condvar=0,
      meany=Mean[[i]];
      meanz=Table[{Mean[[Evidence[[j]]]]},{j,1,k}];
      vary=Var[[i]][[i]];
      If[k==0,
        condmean=Together[meany];
        condvar=Together[vary],
        varz=Table[Table[
                Var[[Evidence[[t]],Evidence[[j]]]],
            {t,1,k}],{j,1,k}];
        covaryz=Table[{Var[[Evidence[[t]]]][[i]]},{t,1,k}];
        zaux=Table[{EviValue[[t]]},{t,1,k}];
        aux=Inverse[varz];
        condmean=meany+Transpose[covaryz].aux.(zaux-meanz);
        condvar=vary-Transpose[covaryz].aux.covaryz;
      ]
    ];
    Print["Evidential nodes =",k," Node =",i];
    Print["Mean =",Together[condmean],
          "Var =",Together[condvar]];
  ]
]
```

Figure 5: A Mathematica program for symbolic propagation of evidence in a Bayesian network.

Known Variables	Damage of the beam X_1	
	Conditional Mean	Conditional Var.
None	0	11.561
$X_9 = 1.0$	0.7	11.071
$X_{10} = 1.0$	$\dfrac{c - cm + 0.7v}{v}$	$\dfrac{-c^2 + 11.071v}{v}$
$X_{11} = x_{11}$	$\dfrac{c - cm + 0.7v + 0.5vx_{11}}{v}$	$\dfrac{-c^2 + 10.821v}{v}$
$X_{12} = 1.0$	$\dfrac{c - cm + 1.2v + 0.5vx_{11}}{v}$	$\dfrac{-c^2 + 10.571v}{v}$
$X_{13} = x_{13}$	$\dfrac{c - cm + 1.2v + 0.5vx_{11} + 1.05vx_{13}}{v}$	$\dfrac{-c^2 + 9.469v}{v}$
$X_{14} = 1.0$	$\dfrac{c - cm + 1.62v + 0.5vx_{11} + 1.05vx_{13}}{v}$	$\dfrac{-c^2 + 9.292v}{v}$

Table 3: Conditional means and variances of X_1, initially and after cumulative evidence.

6 Algebraic Structure of Probabilities

In this section we discuss the algebraic structure of probabilities of single nodes or sets of nodes. We have the following theorem.

THEOREM 2 *The conditional probability distribution of any variable X_i in a Gaussian Bayesian network given any set of other variables in the network is normal with mean and variance which are rational functions, that is, quotients of polynomials, of the evidence variables and the mean and variance or covariance parameters of the initial normal JPD function. The polynomials involved are at the most of degree one in the conditioning variables and in the mean and variance parameters and are of degree two in the covariance parameters. Finally, the polynomial in the denominator is the same for all nodes and for the conditional mean and variance.*

Proof. The proof of the theorem is based on Theorem 1. From (5) we know that the conditional expectation is the sum of $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_{yz}\boldsymbol{\Sigma}_z^{-1}(\mathbf{Z}-\boldsymbol{\mu}_z)$. The last summand is a rational function because we can write it as the quotient of the polynomials $\boldsymbol{\Sigma}_{yz}adj(\boldsymbol{\Sigma}_z)(\mathbf{Z}-\boldsymbol{\mu}_z)$ and $|\boldsymbol{\Sigma}_z|$, where $adj(\boldsymbol{\Sigma}_z)$ is the adjoint matrix of $\boldsymbol{\Sigma}_z$. This implies a rational form of the sum with polynomial denominator $|\boldsymbol{\Sigma}_z|$. Note also that each parameter appears only in one of the three factors of the product $\boldsymbol{\Sigma}_{yz}adj(\boldsymbol{\Sigma}_z)(\mathbf{Z}-\boldsymbol{\mu}_z)$, which implies linearity in each parameter.

Similarly, from (6) we know that the conditional expectation is the sum of $\boldsymbol{\Sigma}_y$ and $-\boldsymbol{\Sigma}_{yz}\boldsymbol{\Sigma}_z^{-1}\boldsymbol{\Sigma}_{yz}^T$. The last summand is a rational function because we can write it as the quotient of the polynomials $-\boldsymbol{\Sigma}_{yz}adj(\boldsymbol{\Sigma}_z)\boldsymbol{\Sigma}_{yz}^T$ and $|\boldsymbol{\Sigma}_z|$. This implies a rational form of the sum with polynomial denominator $|\boldsymbol{\Sigma}_z|$. Note also that all parameters except those in $\boldsymbol{\Sigma}_{yz}$ appear only in one of the factors of the product $-\boldsymbol{\Sigma}_{yz}adj(\boldsymbol{\Sigma}_z)\boldsymbol{\Sigma}_{yz}^T$, which implies linearity in those parameters. On the contrary, the parameters in $\boldsymbol{\Sigma}_{yz}$ appear in two factors and hence they can generate second degree terms in the polynomials.

Finally, we mention that the denominator polynomial can be a second degree in the covariance parameters because of the symmetry of the variance-covariance matrix.

Note that because the denominator polynomial is identical for all possible conditional probabilities with the same evidence, for implementation purposes, it is more convenient to calculate and store all the numerator polynomials for each node and calculate and store the common denominator polynomial separately.

The analysis of the parametric structure of the probabilities in discrete Bayesian networks have shown to be very useful to obtain symbolic results from numeric procedures[7–8]. Analogous results could be obtained for the continuous case using the result given by Theorem 2.

7 Conclusions

Gaussian Bayesian network models have been demonstrated to be useful models to reproduce Engineering problems where dependencies among variables are important factors to be considered. When variables are continuous but limited in range, we can perform a change of variable to transform the new range to the whole real line or choose adequate variances for the values outside the range to have associated small probability. Gaussian Bayesian network models are also very useful to perform a sensitivity analysis using symbolic computations. The conditional means and variances of the nodes given the evidence are shown to be rational functions of the parameters. This parametric structure can be efficiently used in any sensitivity analysis.

Acknowledgments

The authors are grateful to the University of Cantabria, the Dirección General de Investigación Científica y Técnica (DGICYT) (Project PB94-1056), Iberdrola, and NATO Research Office for partial support of this work.

References

1. Bouckaert, R. R., Castillo, E. and Gutiérrez, J. M. (1996), *A Modified Simulation Scheme for Inference in Bayesian Networks*, International Journal of Approximate Reasoning, **4**(1), 55−80.

2. Castillo, E., Cobo, A., Gutiérrez, J.M., Iglesias, A. and Sagástegui, H. (1994), *Causal Network Models in Expert Systems*, Microcomputers in Civil Engineering, **9**, Special issue on “Uncertainty in Expert Systems”, 315−328.

3. Jensen, F. V., Olesen, K. G., and Andersen, S. K. (1990), *An Algebra of Bayesian Belief Universes for Knowledge-Based Systems*, *Networks*, **20**, 637−659.

4. Lauritzen, S. L. and Spiegelhalter, D. J. (1988), *Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems*, Journal of the Royal Statistical Society (B), **50**, 157−224.

5. Pearl, J. (1986), *Fusion, Propagation and Structuring in Belief Networks*, Artificial Intelligence, **29**, 241−288.

6. Pearl, J. (1986), *Evidential Reasoning Using Stochastic Simulation of Causal Models*, Artificial Intelligence, **32**, 245−287.

7. Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1995), *Parametric Structure of Probabilities in Bayesian Networks*, Lectures Notes in Artificial Intelligence, Springer-Verlag, **946**, 89−98.

8. Castillo, E., Gutiérrez, J.M. and Hadi, A.S. (1996), *A New Method for Efficient Symbolic Propagation in Discrete Bayesian Networks*, Networks, In press.

9. Li, Z. and D'Ambrosio, B. (1994), *Efficient Inference in Bayes Nets as a Combinatorial Optimization Problem*, International Journal of Approximate Reasoning, **11**(1), 55−81.

10. Shachter, R. D., D'Ambrosio, B., and DelFabero, B. (1990), *Symbolic Probabilistic Inference in Belief Networks*, in "Proceedings Eighth National Conference on AI", 126−131.

11. Chang K-C. and Fung, R. (1991), *Symbolic Probabilistic Inference with Continuous Variables*, in "Uncertainty in Artificial Intelligence: Proceedings of the Seventh Conference", Morgan Kaufmann, 77−85.

12. Chang, K-C. and Fung, R. (1995), *Symbolic Probabilistic Inference with Both Discrete and Continuous Variables*, IEEE Transactions on Systems, Man and Cybernetics, 25, 6, 910−916.

13. Liu, X., and Li, Z. (1994), *A Reasoning Method in Damage Assessment of Buildings*, Microcomputers in Civil Engineering, **9**, Special issue on "Uncertainty in Expert Systems", **−**.

14. Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1995), *Modelling Probabilistic Networks of Discrete and Continuous Variables*, Technical Report 95−11, Statistics Center, Cornell University.

15. Castillo, E., Gutiérrez, J. M., and Hadi, A. S. (1996), "Expert Systems and Probabilistic Network Models", Springer-Verlag, New York.

16. J. Pearl (1988), "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann, San Mateo, CA.

17. DeGroot (1970), "Optimal Statistical Decisions", McGraw-Hill, New York.

18. Shachter, R. D., and Kenley, C. R. (1989), *Gaussian Influence Diagrams*, Management Science, **35**(5), 527−550.

19. Anderson, T.W. (1984), "An Introduction to Multivariate Analysis," 2^{nd} ed., Prentice Hall, Englewood Cliffs, N.J.

20. Wolfram, S. (1991), "Mathematica: A System for Doing Mathematics by Computer," Addison-Wesley.