Inference in hybrid Bayesian networks with mixtures of truncated basis functions

Helge Langseth[1], Thomas D. Nielsen[2], Rafael Rumí[3], Antonio Salmerón[3]

[1] Department of Computer and Information Science, NTNU (Norway)
[2] Department of Computer Science, Aalborg University (Denmark)
[3] Department of Statistics & Applied Mathematics, University of Almería (Spain)

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
UNIÓN EUROPEA
Fondo Europeo de Desarrollo Regional
"Una manera de hacer Europa"

PGM 2012. Granada, 21 September 2012

Introduction

- MoTBFs provide a flexible framework for hybrid BNs.
- Accurate approximation of known models.
- Learning from data.
- Inference?

Mixtures of truncated basis functions (MoTBFs)

- In what concerns inference in Bayesian networks, initially we only have two types of MoTBFs: univariate and conditional.
- Any other potential showing up during inference is the result of operating over them, namely applying marginalisation and combination.

Mixtures of truncated basis functions (MoTBFs)

Definition (Univariate MoTBF density)

An unconditional MoTBF density over X is:

$$f(x) = \sum_{i=0}^{m-1} a_i \psi_i(x) \quad x \in \Omega_X,$$

where

$$\psi = \{\psi_0(x), \ldots, \psi_{m-1}(x)\}$$

is the set of basis functions for X.

Particular cases

- MTEs: $\Psi = \{1, \exp(-x), \exp(x), \exp(-2x), \exp(2x), \ldots\}$.
- MOPs: $\Psi = \{x^i, i = 0, 1, \ldots\}$.

Mixtures of truncated basis functions (MoTBFs)

Definition (Univariate MoTBF density)

An unconditional MoTBF density over X is:

$$f(x) = \sum_{i=0}^{m-1} a_i \psi_i(x) \quad x \in \Omega_X,$$

where

$$\psi = \{\psi_0(x), \ldots, \psi_{m-1}(x)\}$$

is the set of basis functions for X.

Particular cases

- MTEs: $\Psi = \{1, \exp(-x), \exp(x), \exp(-2x), \exp(2x), \ldots\}$.
- MOPs: $\Psi = \{x^i, i = 0, 1, \ldots\}$.

Mixtures of truncated basis functions (MoTBFs)

H. Langseth, T.D. Nielsen, R. Rumí, A. Salmerón

Definition (Conditional MoTBF density)

- $\mathbf{Y}$ and $\mathbf{Z}$: discrete and continuous variables.
- Domain of $\mathbf{Z}$ divided into k hypercubes.
- For each $\mathbf{y}$, and each hypercube $j = 1, \ldots, k$, the conditional MoTBF density of X given $\mathbf{Z}$ and $\mathbf{Y}$ is

$$f(x|\mathbf{z}, \mathbf{y}) = \sum_{i=0}^{m-1} a_{i,j} \psi_i(x)$$

Mixtures of truncated basis functions (MoTBFs)

H. Langseth,
T.D. Nielsen,
R. Rumí,
A. Salmerón

Potential advantages of MoTBFs for inference

- Univariate MoTBFs do not require domain splitting (unlike classical approach to MTEs and MOPs).
- Conditional MoTBFs are piecewise univariate over the head variable.
- As a consequence, each variable in the BN explicitly appears in only one potential initially.
- If a variable appears in a potential not as a head variable, it only determines the hypercubes of the conditional density.
- One can consider a fixed set of possible split points for each variable, regardless of the function where it appears.
- There is no need to explicitly store the basis functions.

Combination

H. Langseth,
T.D. Nielsen,
R. Rumí,
A. Salmerón

$$f(x_1) = \sum_{i=0}^{m-1} a^1_{i,h}\psi_i(x_1) \quad ; \quad f(x_2) = \sum_{i=0}^{m-1} a^2_{i,t}\psi_i(x_2)$$

$$f(x_1, x_2) = \left(\sum_{i=0}^{m-1} a^1_{i,h}\psi_i(x_1)\right)\left(\sum_{i=0}^{m-1} a^2_{i,t}\psi_i(x_2)\right)$$

Combination

$$f(x_1) = \sum_{i=0}^{m-1} a^1_{i,h}\psi_i(x_1) \quad ; \quad f(x_2) = \sum_{i=0}^{m-1} a^2_{i,t}\psi_i(x_2)$$

$$f(x_1, x_2) = \left(\sum_{i=0}^{m-1} a^1_{i,h}\psi_i(x_1)\right)\left(\sum_{i=0}^{m-1} a^2_{i,t}\psi_i(x_2)\right)$$

MoTBF potential

Combination

$$f(x_1) = \sum_{i=0}^{m-1} a_{i,h}^1 \psi_i(x_1) \quad ; \quad f(x_2) = \sum_{i=0}^{m-1} a_{i,t}^2 \psi_i(x_2)$$

$$f(x_1, x_2) = \left(\sum_{i=0}^{m-1} a_{i,h}^1 \psi_i(x_1)\right)\left(\sum_{i=0}^{m-1} a_{i,t}^2 \psi_i(x_2)\right)$$

MoTBF potential MoTBF potential

Combination

$$f(x_1) = \sum_{i=0}^{m-1} a_{i,h}^1 \psi_i(x_1) \quad ; \quad f(x_2) = \sum_{i=0}^{m-1} a_{i,t}^2 \psi_i(x_2)$$

Factorised MoTBF potential

$$f(x_1, x_2) = \left(\sum_{i=0}^{m-1} a_{i,h}^1 \psi_i(x_1)\right)\left(\sum_{i=0}^{m-1} a_{i,t}^2 \psi_i(x_2)\right)$$

MoTBF potential

MoTBF potential

Marginalisation

H. Langseth, T.D. Nielsen, R. Rumí, A. Salmerón

Proposition

$$f_{\mathbf{Z}\setminus\{Z_j\}}(z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_c) =$$
$$\sum_{l=1}^{r}\left(\prod_{i\neq j}\sum_{s=0}^{m-1} a^i_{s,\cdot,(h,l)}\psi_s(z_i)\right)\left(\sum_{s=0}^{m-1} a^j_{s,\cdot,(h,l)}\int_{\Omega^l_{Z_j}}\psi_s(z_j)dz_j\right).$$

- Bad news: NOT a factorised MoTBF potential!
- Good news: the integrals can be computed off-line, prior to inference, if the split points are fixed.

Marginalisation

Proposition

$$f_{\mathbf{Z}\setminus\{Z_j\}}(z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_c) = \sum_{l=1}^{r}\left(\prod_{i\neq j}\sum_{s=0}^{m-1} a^i_{s,\cdot,(h,l)}\psi_s(z_i)\right)\left(\sum_{s=0}^{m-1} a^j_{s,\cdot,(h,l)}\int_{\Omega^l_{Z_j}}\psi_s(z_j)dz_j\right).$$

- Bad news: NOT a factorised MoTBF potential!
- Good news: the integrals can be computed off-line, prior to inference, if the split points are fixed.

This kind of potential is called SP factorised MoTBF potential

Combination of SP factorised potentials

$$f_{\mathbf{X}_1}(\mathbf{x}_1) = \sum_{l=1}^{r_1} f^l_{\mathbf{X}_1}(\mathbf{x}_1) \ ; \ f_{\mathbf{X}_2}(\mathbf{x}_2) = \sum_{m=1}^{r_2} f^m_{\mathbf{X}_2}(\mathbf{x}_2)$$

The combination of $f_{\mathbf{X}_1}$ and $f_{\mathbf{X}_2}$ is a new potential over variables $\mathbf{X}_{12} = \mathbf{X}_1 \cup \mathbf{X}_2$ defined as

$$f(\mathbf{x}_{12}) = \sum_{l=1}^{r_1} \sum_{m=1}^{r_2} f^l_{\mathbf{X}_1}(\mathbf{x}_{12}^{\downarrow \mathbf{X}_1}) f^m_{\mathbf{X}_2}(\mathbf{x}_{12}^{\downarrow \mathbf{X}_2}),$$

which is an SP factorised potential.

Marginalisation of SP factorised potentials

$$f_{\mathbf{Z}}(\mathbf{z}) = \sum_{l=1}^{r} f^{l}_{\mathbf{Z}}(\mathbf{z})$$

$$f_{\mathbf{Z}\setminus\{Z_j\}}(z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_c) = \sum_{l=1}^{r} f^{l}_{\mathbf{Z}\setminus\{Z_j\}}(z_1,\ldots,z_{j-1},z_{j+1},\ldots,z_n)$$

Again, it is an SP factorised potential.

Marginalisation of SP factorised potentials

$$f_{\mathbf{Z}}(\mathbf{z}) = \sum_{l=1}^{r} f_{\mathbf{Z}}^{l}(\mathbf{z})$$

$$f_{\mathbf{Z}\setminus\{Z_j\}}(z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_c) = \sum_{l=1}^{r} f_{\mathbf{Z}\setminus\{Z_j\}}^{l}(z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n)$$

Again, it is an SP factorised potential.

Why are SP factorised potentials of interest?

- They are closed for marginalisation and combination.
- Hence, inference algorithms as Shenoy-Shafer and Variable Elimination can be used.
- Operations over them are lazy by nature, i.e., handling them actually consists of handling sets of function (storing, indexing and retrieving them).

Classical MTE calculations vs. MoTBFs

Two experiments conducted

1. MoTBF vs. classical MTE approach:
   - No splits in head variables.
   - Fixed splits in conditionals.
2. Lazy operations on SP factorised potentials vs. classical MTE operations.
   - Random split points everywhere.

We use the Variable Elimination algorithm.

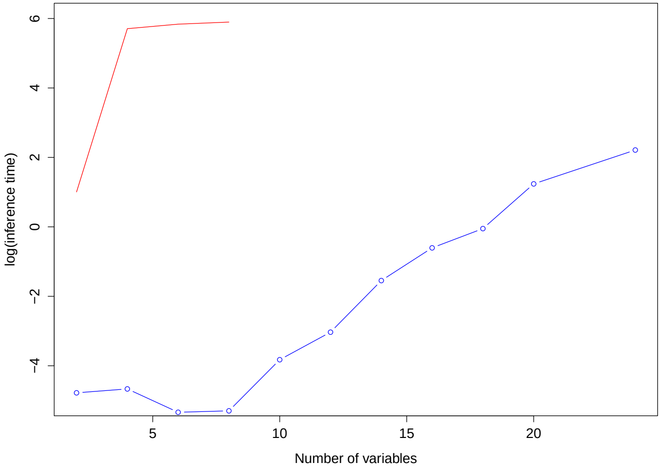
Classical MTE calculations vs. MoTBFs

Two experiments conducted

1. MoTBF vs. classical MTE approach:
   - No splits in head variables.
   - Fixed splits in conditionals.
2. Lazy operations on SP factorised potentials vs. classical MTE operations.
   - Random split points everywhere.

We use the Variable Elimination algorithm.

Experimental results



Experimental results

- Run time order of magnitude lower for the new approach.
- SP factorised potentials (lazy operations) by themselves, provide significant improvements.
- Storage requirements clearly lower.
- See the poster for the detailed numbers.

Experimental results

- Run time order of magnitude lower for the new approach.
- SP factorised potentials (lazy operations) by themselves, provide significant improvements.
- Storage requirements clearly lower.
- See the poster for the detailed numbers.

Conclusions

- MoTBF framework provides important advantages for inference (No split points in head variables, fixed set of basis functions, lazy operations).
- Lazy operations can be used (SP factorised potentials) even under classical MTE calculations.
- Significant advance wrt. previous MTE calculations: savings in space requirements and increase in efficiency.

Future work

- Explore the impact (benefit) in efficiency of the use of fixed split points and its possible tradeoff with accuracy.
- Experiments including evidences and incorporating discrete variables.
- Re-approximate large potentials.
- Analyse other benefits of using basis functions for inference.