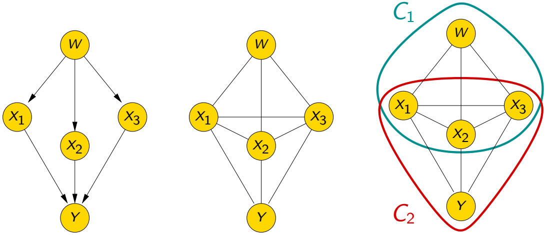
Tensor rank-one decomposition of probability tables

Petr Savický and Jiří Vomlel

Academy of Sciences of the Czech Republic (AV ČR)

Paris, June 6, 2006

Probabilistic inference with Bayesian networks (the junction tree method)

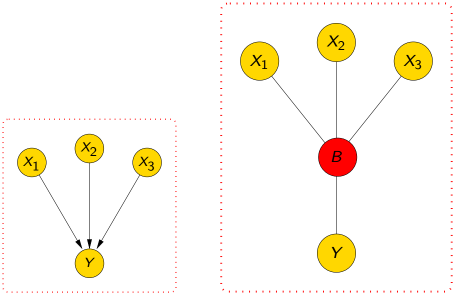


$$2^4 + 2^4 = 32$$

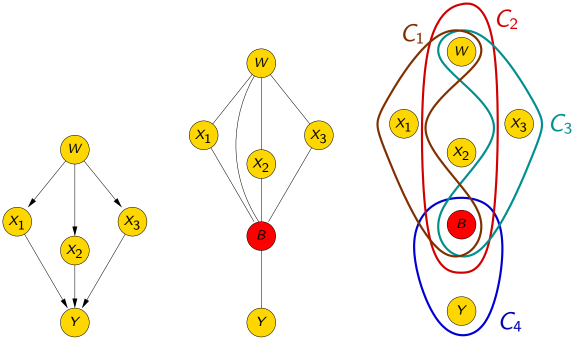
Decomposition of a conditional probability table

We say that $\psi : \mathcal{X}_1 \times \ldots \times \mathcal{X}_m \times \mathcal{Y} \mapsto \mathbb{R}$ can be factorized by use of variable B if there exist potentials $\xi : \mathcal{Y} \times \mathcal{B} \mapsto \mathbb{R}$ and $\varphi_i : \mathcal{X}_i \times \mathcal{B} \mapsto \mathbb{R}$, for $i = 1, \ldots, m$ such that for all $(x_1, \ldots, x_m, y) \in \mathcal{X}_1 \times \ldots \mathcal{X}_m \times \mathcal{Y}$, we have

$$\psi(x_1, \ldots, x_m, y) = \sum_{b \in \mathcal{B}} \xi(y, b) \cdot \prod_{i=1}^{m} \varphi_i(x_i, b)$$

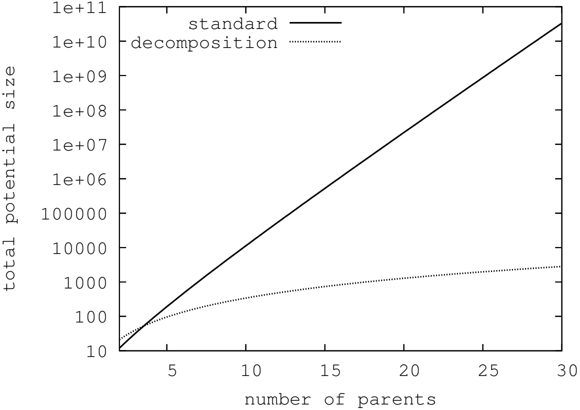


Probabilistic inference after the decomposition



If B has two states:
$2^3 + 2^3 + 2^3 + 2^2 = 28$
which is less than 32.

The savings with respect to number of parents



Comparison of the total size of junction tree.

Correspondence to tensor rank-one decomposition

A decomposition using the auxiliary variable B

$$\psi(x_1, \ldots, x_m, y) \quad = \quad \sum_{b \in \mathcal{B}} \xi(y, b) \cdot \prod_{i=1}^{m} \varphi_i(x_i, b)$$

that has the minimal number of states of B

is in fact a (minimal) tensor rank-one decomposition of tensor ψ.

Correspondence to tensor rank-one decomposition

A decomposition using the auxiliary variable B

$$\psi(x_1, \ldots, x_m, y) \; = \; \sum_{b \in \mathcal{B}} \xi(y, b) \cdot \prod_{i=1}^{m} \varphi_i(x_i, b)$$

that has the minimal number of states of B

is in fact a (minimal) tensor rank-one decomposition of tensor ψ.

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \quad = \quad (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) \quad = \quad \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \;=\; \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \;=\; (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) \;=\; \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) \;=\; \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \quad = \quad (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) \quad = \quad \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \quad = \quad (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) \quad = \quad \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \quad = \quad (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) \quad = \quad \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) = \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} = (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\psi(x_1, \ldots, x_n) = \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n)$$

$$\varphi^b(x_1, \ldots, x_n) = \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b)$$

Tensor rank

Let $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_n$. A n-th order tensor $\varphi : \mathcal{X} \mapsto \mathbb{R}$ has rank 1 if it is the outer product of n vectors:

$$\varphi(x_1, \ldots, x_n) \quad = \quad \varphi_1(x_1) \cdot \ldots \cdot \varphi_n(x_n)$$

An example of a 3rd-order tensor that has rank 1:

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} \quad = \quad (1,2) \otimes (1,2) \otimes (1,2)$$

Rank r of a tensor ψ is the minimal number of rank-one tensors necessary to add in order to yield ψ.

$$\begin{aligned} \psi(x_1, \ldots, x_n) &= \sum_{b=1}^{r} \varphi^b(x_1, \ldots, x_n) \\ \varphi^b(x_1, \ldots, x_n) &= \varphi_1(x_1, b) \cdot \ldots \cdot \varphi_n(x_n, b) \end{aligned}$$

Example of a rank-2 tensor

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,9) \end{pmatrix} = \begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} + \begin{pmatrix} (0,0) & (0,0) \\ (0,0) & (0,1) \end{pmatrix}$$

$$= (1,2) \otimes (1,2) \otimes (1,2) + (0,1) \otimes (0,1) \otimes (0,1)$$

$$\neq (a,b) \otimes (c,d) \otimes (e,f) \text{ for any } a,b,c,d,e,f \in \mathbb{R}.$$

Example of a rank-2 tensor

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,9) \end{pmatrix} = \begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} + \begin{pmatrix} (0,0) & (0,0) \\ (0,0) & (0,1) \end{pmatrix}$$

$$= (1,2) \otimes (1,2) \otimes (1,2) + (0,1) \otimes (0,1) \otimes (0,1)$$

$$\neq (a,b) \otimes (c,d) \otimes (e,f) \text{ for any } a,b,c,d,e,f \in \mathbb{R}.$$

Example of a rank-2 tensor

$$\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,9) \end{pmatrix} = \begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} + \begin{pmatrix} (0,0) & (0,0) \\ (0,0) & (0,1) \end{pmatrix}$$

$$= (1,2) \otimes (1,2) \otimes (1,2) \; + \; (0,1) \otimes (0,1) \otimes (0,1)$$

$$\neq (a,b) \otimes (c,d) \otimes (e,f) \text{ for any } a,b,c,d,e,f \in \mathbb{R}.$$

Example of a rank-2 tensor

$$\begin{aligned}
\begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,9) \end{pmatrix} &= \begin{pmatrix} (1,2) & (2,4) \\ (2,4) & (4,8) \end{pmatrix} + \begin{pmatrix} (0,0) & (0,0) \\ (0,0) & (0,1) \end{pmatrix} \\
&= (1,2) \otimes (1,2) \otimes (1,2) \; + \; (0,1) \otimes (0,1) \otimes (0,1) \\
&\neq (a,b) \otimes (c,d) \otimes (e,f) \;\; \text{for any } a,b,c,d,e,f \in \mathbb{R}.
\end{aligned}$$

Tensors of our special interest

Definition (Indicator function)

$$I(expr) \quad = \quad \begin{cases} 1 & \text{if } expr \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

Definition (Tensor ψ_f that represent a functional dependence)

Given a function $f : \mathcal{X} \mapsto \mathcal{Y}$ the tensor defined for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ as

$$\psi_f(x, y) = I(y = f(x))$$

is a tensor ψ_f that represent a functional dependence of one variable Y on variables $X_1, \ldots, X_m$.

Tensors of our special interest

Definition (Indicator function)

$$I(expr) = \begin{cases} 1 & \text{if } expr \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

Definition (Tensor ψ_f that represent a functional dependence)

Given a function $f : \mathcal{X} \mapsto \mathcal{Y}$ the tensor defined for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ as

$$\psi_f(x, y) = I(y = f(x))$$

is a tensor ψ_f that represent a functional dependence of one variable Y on variables $X_1, \ldots, X_m$.

Example of a tensor representing functional dependence

- $\mathcal{X}_i = \{0, 1\}$ for $i = 1, 2$,
- $f(x_1, x_2) = x_1 + x_2$, and
- $\mathcal{Y} = \{0, 1, 2\}$.

$$
\begin{aligned}
\psi_f(x_1, x_2, y) &= I(y = x_1 + x_2) \\
&= \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \end{pmatrix}
\end{aligned}
$$

Example of a tensor representing functional dependence

- $\mathcal{X}_i = \{0, 1\}$ for $i = 1, 2$,
- $f(x_1, x_2) = x_1 + x_2$, and
- $\mathcal{Y} = \{0, 1, 2\}$.

$$\begin{aligned} \psi_f(x_1, x_2, y) &= I(y = x_1 + x_2) \\ &= \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \end{aligned}$$

Results

- $\mathcal{X}_i = [0, a_i]$ for $i = 1, \ldots, m$
- function $f : \mathcal{X} \mapsto \mathcal{Y}$ and
- $\psi_f : \mathcal{X} \times \mathcal{Y} \mapsto \{0, 1\}$ be a tensor representing the functional dependence given by f.

Lemma (Rank lower bound)

$rank(\psi_f) \geq |\mathcal{Y}|$.

Lemma (Minimimum, maximum)

If $f(x) = \min\{x_1, \ldots, x_m\}$ *or* $\max\{x_1, \ldots, x_m\}$ *then* $rank(\psi_f) = |\mathcal{Y}|$.

Lemma (Addition)

If $f(x) = \sum_{i=1}^{m} x_i$ *then* $rank(\psi_f) = |\mathcal{Y}|$.

Results

- $\mathcal{X}_i = [0, a_i]$ for $i = 1, \ldots, m$
- function $f : \mathcal{X} \mapsto \mathcal{Y}$ and
- $\psi_f : \mathcal{X} \times \mathcal{Y} \mapsto \{0, 1\}$ be a tensor representing the functional dependence given by f.

Lemma (Rank lower bound)

$rank(\psi_f) \geq |\mathcal{Y}|$.

Lemma (Minimimum, maximum)

If $f(x) = \min\{x_1, \ldots, x_m\}$ or $\max\{x_1, \ldots, x_m\}$ then $rank(\psi_f) = |\mathcal{Y}|$.

Lemma (Addition)

If $f(x) = \sum_{i=1}^{m} x_i$ then $rank(\psi_f) = |\mathcal{Y}|$.

Results

- $\mathcal{X}_i = [0, a_i]$ for $i = 1, \ldots, m$
- function $f : \mathcal{X} \mapsto \mathcal{Y}$ and
- $\psi_f : \mathcal{X} \times \mathcal{Y} \mapsto \{0, 1\}$ be a tensor representing the functional dependence given by f.

Lemma (Rank lower bound)

$rank(\psi_f) \geq |\mathcal{Y}|$.

Lemma (Minimimum, maximum)

If $f(x) = \min\{x_1, \ldots, x_m\}$ *or* $\max\{x_1, \ldots, x_m\}$ *then* $rank(\psi_f) = |\mathcal{Y}|$.

Lemma (Addition)

If $f(x) = \sum_{i=1}^{m} x_i$ *then* $rank(\psi_f) = |\mathcal{Y}|$.

Results

- $\mathcal{X}_i = [0, a_i]$ for $i = 1, \ldots, m$
- function $f : \mathcal{X} \mapsto \mathcal{Y}$ and
- $\psi_f : \mathcal{X} \times \mathcal{Y} \mapsto \{0, 1\}$ be a tensor representing the functional dependence given by f.

Lemma (Rank lower bound)

$rank(\psi_f) \geq |\mathcal{Y}|$.

Lemma (Minimimum, maximum)

If $f(x) = \min\{x_1, \ldots, x_m\}$ or $\max\{x_1, \ldots, x_m\}$ then $rank(\psi_f) = |\mathcal{Y}|$.

Lemma (Addition)

If $f(x) = \sum_{i=1}^{m} x_i$ then $rank(\psi_f) = |\mathcal{Y}|$.

Conclusions

- Standard methods for inference in Bayesian networks may become intractable if a variable has many parents.
- Tensor rank-one decomposition open new possibilities for more compact inference structures.
- We have shown how (minimal) tensor rank-one decompositions can be found for some special tensors representing functional dependence.
- These results can be easily generalized to noisy functional dependence.

Conclusions

- Standard methods for inference in Bayesian networks may become intractable if a variable has many parents.
- Tensor rank-one decomposition open new possibilities for more compact inference structures.
- We have shown how (minimal) tensor rank-one decompositions can be found for some special tensors representing functional dependence.
- These results can be easily generalized to noisy functional dependence.

Conclusions

- Standard methods for inference in Bayesian networks may become intractable if a variable has many parents.
- Tensor rank-one decomposition open new possibilities for more compact inference structures.
- We have shown how (minimal) tensor rank-one decompositions can be found for some special tensors representing functional dependence.
- These results can be easily generalized to noisy functional dependence.

Conclusions

- Standard methods for inference in Bayesian networks may become intractable if a variable has many parents.
- Tensor rank-one decomposition open new possibilities for more compact inference structures.
- We have shown how (minimal) tensor rank-one decompositions can be found for some special tensors representing functional dependence.
- These results can be easily generalized to noisy functional dependence.