Additive decomposition of probability tables

Petr Savický and Jiří Vomlel

Academy of Sciences of the Czech Republic (AV ČR)

Třešť, December 12, 2006

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Random variables

Binary random variables $X_1, X_2, X_3, X_4, X_5, X_6$:

- X_1 ... coronary hearth disease (states: 0/1)
- X_2 ... high systolic blood pressure (states: 0/1)
- X_3 ... high diastolic blood pressure (states: 0/1)
- X_4 ... high cholesterol (states: 0/1)
- X_5 ... physical activity (states: 0/1)
- X_6 ... family anamnesis (states: 0/1)

Discrete probability distribution

Discrete probability distribution $P(X_1, X_2, X_3, X_4, X_5, X_6)$ can be represented by a table:

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0	0.0342	0.0315	0.0319	0.0306	0.0314	0.0328	0.0343	0.0325
		1	0.0284	0.0295	0.0307	0.0289	0.0318	0.0302	0.0319	0.0288
	1	0	0.0287	0.0305	0.0332	0.0282	0.0308	0.0319	0.0311	0.0295
		1	0.0286	0.0296	0.0276	0.0317	0.0304	0.0256	0.0282	0.0252
1	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0001
		1	0.0001	0.0001	0.0001	0.0000	0.0000	0.0005	0.0002	0.0004
	1	0	0.0000	0.0004	0.0002	0.0002	0.0006	0.0010	0.0006	0.0020
		1	0.0003	0.0014	0.0009	0.0016	0.0029	0.0063	0.0029	0.0068

Discrete probability distribution

Discrete probability distribution $P(X_1, X_2, X_3, X_4, X_5, X_6)$ can be represented by a table:

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0	0.0342	0.0315	0.0319	0.0306	0.0314	0.0328	0.0343	0.0325
		1	0.0284	0.0295	0.0307	0.0289	0.0318	0.0302	0.0319	0.0288
	1	0	0.0287	0.0305	0.0332	0.0282	0.0308	0.0319	0.0311	0.0295
		1	0.0286	0.0296	0.0276	0.0317	0.0304	0.0256	0.0282	0.0252
1	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0001
		1	0.0001	0.0001	0.0001	0.0000	0.0000	0.0005	0.0002	0.0004
	1	0	0.0000	0.0004	0.0002	0.0002	0.0006	0.0010	0.0006	0.0020
		1	0.0003	0.0014	0.0009	0.0016	0.0029	0.0063	0.0029	0.0068

Querries

Conditional probability $P(X_1 \mid X_6 = 0)$,
i.e., probability of coronary hearth disease given negative family anamnesis
Bayes rule

$$P(X_1 \mid X_6 = 0) = \frac{P(X_1, X_6 = 0)}{P(X_6 = 0)} = \frac{P(X_1, X_6 = 0)}{\sum_{x_1=0}^{1} P(X_1 = x_1, X_6 = 0)}$$

$\sum_{x_i=0}^{1} P(\ldots, X_i = x_i, \ldots)$ denotes marginalizing out variable X_i.

Querries

Conditional probability $P(X_1 \mid X_6 = 0)$,
i.e., probability of coronary hearth disease given negative family anamnesis

Bayes rule

$$P(X_1 \mid X_6 = 0) = \frac{P(X_1, X_6 = 0)}{P(X_6 = 0)} = \frac{P(X_1, X_6 = 0)}{\sum_{x_1=0}^{1} P(X_1 = x_1, X_6 = 0)}$$

$\sum_{x_i=0}^{1} P(\ldots, X_i = x_i, \ldots)$ denotes marginalizing out variable X_i.

Querries

Conditional probability $P(X_1 \mid X_6 = 0)$,
i.e., probability of coronary hearth disease given negative family anamnesis
Bayes rule

$$P(X_1 \mid X_6 = 0) = \frac{P(X_1, X_6 = 0)}{P(X_6 = 0)}$$

$$= \frac{P(X_1, X_6 = 0)}{\sum_{x_1=0}^{1} P(X_1 = x_1, X_6 = 0)}$$

$\sum_{x_i=0}^{1} P(\ldots, X_i = x_i, \ldots)$ denotes marginalizing out variable X_i.

Querries

Conditional probability $P(X_1 \mid X_6 = 0)$,
i.e., probability of coronary hearth disease given negative family anamnesis
Bayes rule

$$
\begin{aligned}
P(X_1 \mid X_6 = 0) &= \frac{P(X_1, X_6 = 0)}{P(X_6 = 0)} \\
&= \frac{P(X_1, X_6 = 0)}{\sum_{x_1=0}^{1} P(X_1 = x_1, X_6 = 0)}
\end{aligned}
$$

$\sum_{x_i=0}^{1} P(\ldots, X_i = x_i, \ldots)$ denotes marginalizing out variable X_i.

Querries

Conditional probability $P(X_1 \mid X_6 = 0)$,
i.e., probability of coronary hearth disease given negative family anamnesis
Bayes rule

$$P(X_1 \mid X_6 = 0) = \frac{P(X_1, X_6 = 0)}{P(X_6 = 0)} = \frac{P(X_1, X_6 = 0)}{\sum_{x_1=0}^{1} P(X_1 = x_1, X_6 = 0)}$$

$\sum_{x_i=0}^{1} P(\ldots, X_i = x_i, \ldots)$ denotes marginalizing out variable X_i.

Computation of $P(X_1, X_6 = 0)$

$$P(X_1, X_6 = 0) = \sum_{x_2=0}^{1} \sum_{x_3=0}^{1} \sum_{x_4=0}^{1} \sum_{x_5=0}^{1} P(X_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = 0)$$

Computation of $P(X_1, X_6 = 0)$

$$P(X_1, X_6 = 0) = \sum_{x_2=0}^{1} \sum_{x_3=0}^{1} \sum_{x_4=0}^{1} \sum_{x_5=0}^{1} P(X_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = 0)$$

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0	0.0342	0.0315	0.0319	0.0306	0.0314	0.0328	0.0343	0.0325
		1	0.0284	0.0295	0.0307	0.0289	0.0318	0.0302	0.0319	0.0288
	1	0	0.0287	0.0305	0.0332	0.0282	0.0308	0.0319	0.0311	0.0295
		1	0.0286	0.0296	0.0276	0.0317	0.0304	0.0256	0.0282	0.0252
1	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.0001
		1	0.0001	0.0001	0.0001	0.0000	0.0000	0.0005	0.0002	0.0004
	1	0	0.0000	0.0004	0.0002	0.0002	0.0006	0.0010	0.0006	0.0020
		1	0.0003	0.0014	0.0009	0.0016	0.0029	0.0063	0.0029	0.0068

Computation of $P(X_1, X_6 = 0)$

$$P(X_1, X_6 = 0) \quad = \quad \sum_{x_2=0}^{1} \sum_{x_3=0}^{1} \sum_{x_4=0}^{1} \sum_{x_5=0}^{1} P(X_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = 0)$$

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0	0.0342	0.0315	0.0319	0.0306				
		1	0.0284	0.0295	0.0307	0.0289				
	1	0	0.0287	0.0305	0.0332	0.0282				
		1	0.0286	0.0296	0.0276	0.0317				
1	0	0	0.0000	0.0000	0.0000	0.0000				
		1	0.0001	0.0001	0.0001	0.0000				
	1	0	0.0000	0.0004	0.0002	0.0002				
		1	0.0003	0.0014	0.0009	0.0016				

Computation of $P(X_1, X_6 = 0)$

$$P(X_1, X_6 = 0) \quad = \quad \sum_{x_2=0}^{1} \sum_{x_3=0}^{1} \sum_{x_4=0}^{1} \sum_{x_5=0}^{1} P(X_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = 0)$$

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0	0.0342	0.0315	0.0319	0.0306				
		1	0.0284	0.0295	0.0307	0.0289				
	1	0	0.0287	0.0305	0.0332	0.0282				
		1	0.0286	0.0296	0.0276	0.0317				
1	0	0	0.0000	0.0000	0.0000	0.0000				
		1	0.0001	0.0001	0.0001	0.0000				
	1	0	0.0000	0.0004	0.0002	0.0002				
		1	0.0003	0.0014	0.0009	0.0016				

Computation of $P(X_1, X_6 = 0)$

$$P(X_1, X_6 = 0) = \sum_{x_2=0}^{1} \sum_{x_3=0}^{1} \sum_{x_4=0}^{1} \sum_{x_5=0}^{1} P(X_1, X_2 = x_2, X_3 = x_3, X_4 = x_4, X_5 = x_5, X_6 = 0)$$

		X_6	0				1			
		X_5	0		1		0		1	
		X_4	0	1	0	1	0	1	0	1
X_1	X_2	X_3								
0	0	0								
		1		0.4838					after 15 additions	
	1	0								
		1								
1	0	0								
		1		0.0053					after 15 additions	
	1	0								
		1								

Computational complexity

- Probability distribution over n binary variables $P(X_1, \ldots, X_n)$.
- Generally, the computation of $P(X_i \mid X_j = x_j)$ requires $2(2^{n-2} - 1)$ additions.
- **It is exponential in number of variables and thus intractable for larger n!**
- If one addition takes $0.001 \mu s$ $(= \frac{1}{1GHz})$ then for $n = 50$ we need 13 days to compute the marginal distribution!

Computational complexity

- Probability distribution over n binary variables $P(X_1, \ldots, X_n)$.
- Generally, the computation of $P(X_i \mid X_j = x_j)$ requires $2(2^{n-2} - 1)$ additions.
- **It is exponential in number of variables and thus intractable for larger n!**
- If one addition takes $0.001\mu s$ $(= \frac{1}{1GHz})$ then for $n = 50$ we need 13 days to compute the marginal distribution!

Computational complexity

- Probability distribution over n binary variables $P(X_1, \ldots, X_n)$.
- Generally, the computation of $P(X_i \mid X_j = x_j)$ requires $2(2^{n-2} - 1)$ additions.
- **It is exponential in number of variables and thus intractable for larger n!**
- If one addition takes $0.001\mu s$ $(= \frac{1}{1GHz})$ then for $n = 50$ we need 13 days to compute the marginal distribution!

Computational complexity

- Probability distribution over n binary variables $P(X_1, \ldots, X_n)$.
- Generally, the computation of $P(X_i \mid X_j = x_j)$ requires $2(2^{n-2} - 1)$ additions.
- **It is exponential in number of variables and thus intractable for larger n!**
- If one addition takes $0.001 \mu s$ $(= \frac{1}{1GHz})$ then for $n = 50$ we need 13 days to compute the marginal distribution!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) = \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$P(X_1, X_n = 0)$$
$$= \sum_{x_2=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0)$$
$$= \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0)$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) = \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$\begin{aligned} P(X_1, X_n = 0) &= \sum_{x_2=0}^{1} \ldots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0) \\ &= \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0) \end{aligned}$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) = \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$\begin{aligned} P(X_1, X_n = 0) \\ &= \sum_{x_2=0}^{1} \ldots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0) \\ &= \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0) \end{aligned}$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) \quad = \quad \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$P(X_1, X_n = 0)$$
$$= \sum_{x_2=0}^{1} \ldots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0)$$
$$= \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0)$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) = \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$\begin{aligned} & P(X_1, X_n = 0) \\ & = \sum_{x_2=0}^{1} \ldots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0) \\ & = \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0) \end{aligned}$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

Can we decrease the complexity?

Yes, if we exploit the internal structure of the probability distribution $P(X_1, \ldots, X_n)$. For example, if

$$P(X_1, \ldots, X_n) \;=\; \psi(X_1) \cdot \ldots \cdot \psi(X_n)$$

then computation of $P(X_1, X_n = 0)$ can be performed as

$$\begin{aligned} & P(X_1, X_n = 0) \\ & = \sum_{x_2=0}^{1} \cdots \sum_{x_{n-1}=0}^{1} \psi(X_1) \cdot \psi(X_2 = x_2) \cdot \ldots \cdot \psi(X_{n-1} = x_{n-1}) \cdot \psi(X_n = 0) \\ & = \psi(X_1) \left(\sum_{x_2=0}^{1} \psi(X_2 = x_2) \right) \cdot \ldots \cdot \left(\sum_{x_{n-1}=0}^{1} \psi(X_{n-1} = x_{n-1}) \right) \psi(X_n = 0) \end{aligned}$$

which requires only $n-2$ additions and $2(n-1)$ multiplications!

What can we do if we are not so lucky?

Not allways distribution $P(X_1, \ldots, X_n)$ has such a nice internal structure as in the previous case.

However, we can exploit also more complicated internal structures. Either by:

(1) multiplicative decomposition or

(2) additive decomposition

What can we do if we are not so lucky?

Not allways distribution $P(X_1, \ldots, X_n)$ has such a nice internal structure as in the previous case.

However, we can exploit also more complicated internal structures. Either by:

(1) multiplicative decomposition or

(2) additive decomposition

What can we do if we are not so lucky?

Not allways distribution $P(X_1, \ldots, X_n)$ has such a nice internal structure as in the previous case.

However, we can exploit also more complicated internal structures. Either by:

(1) multiplicative decomposition or

(2) additive decomposition

What can we do if we are not so lucky?

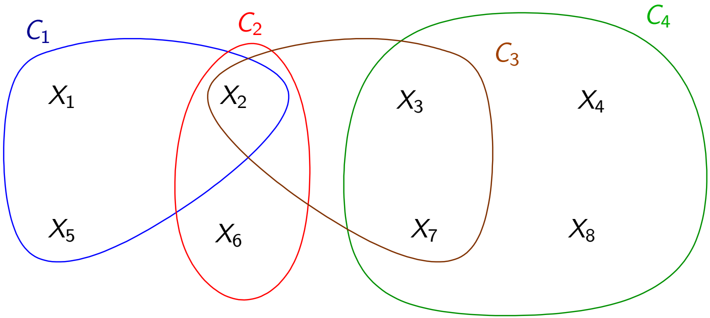
Not allways distribution $P(X_1, \ldots, X_n)$ has such a nice internal structure as in the previous case.

However, we can exploit also more complicated internal structures. Either by:

(1) multiplicative decomposition or

(2) additive decomposition

Multiplicative decomposition



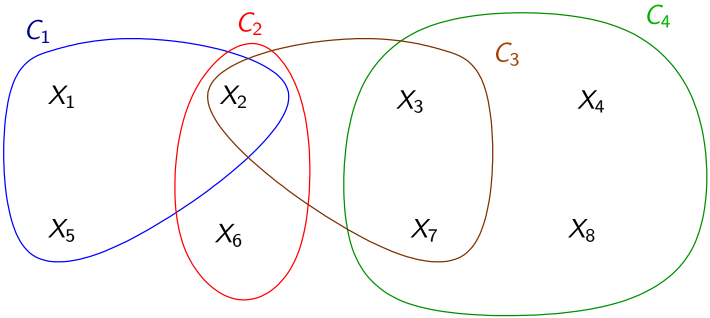
If $C_i \subset \{1, \ldots, n\}, i \in \{1, \ldots, k\}$ are edges of a decomposable hypergraph and

$$P(X_1, \ldots, X_n) = \psi(X_{C_1}) \cdot \ldots \cdot \psi(X_{C_k})$$

then in order to get $P(X_1, X_n = 0)$ we need the number of additions and multiplications proportional to the state space of the largest set C_i, i.e., to

$$2^{|C_i|}$$

Multiplicative decomposition



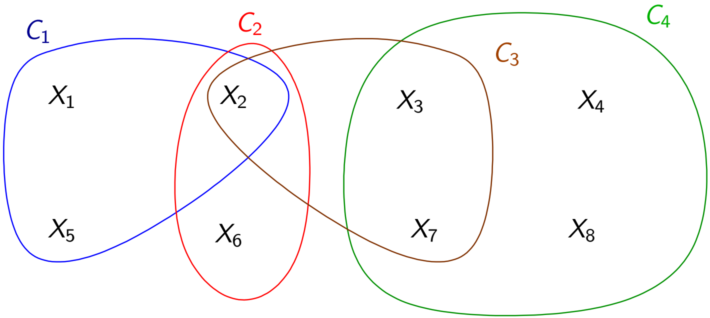
If $C_i \subset \{1, \ldots, n\}, i \in \{1, \ldots, k\}$ are edges of a decomposable hypergraph and

$$P(X_1, \ldots, X_n) \quad = \quad \psi(X_{C_1}) \cdot \ldots \cdot \psi(X_{C_k})$$

then in order to get $P(X_1, X_n = 0)$ we need the number of additions and multiplications proportional to the state space of the largest set C_i, i.e., to

$$2^{|C_i|}$$

Multiplicative decomposition



If $C_i \subset \{1, \ldots, n\}, i \in \{1, \ldots, k\}$ are edges of a decomposable hypergraph and

$$P(X_1, \ldots, X_n) = \psi(X_{C_1}) \cdot \ldots \cdot \psi(X_{C_k})$$

then in order to get $P(X_1, X_n = 0)$ we need the number of additions and multiplications proportional to the state space of the largest set C_i, i.e., to

$$2^{|C_i|}$$

Additive decomposition

$$P(X_1, \ldots, X_n) = \sum_{q=1}^{r} \psi_q(X_1) \cdot \ldots \cdot \psi_q(X_n)$$

In order to get $P(X_1, X_n = 0)$ we need $r(n-1)$ multiplications and $r(n-2)$ additions.

Additive decomposition

$$P(X_1, \ldots, X_n) = \sum_{q=1}^{r} \psi_q(X_1) \cdot \ldots \cdot \psi_q(X_n)$$

In order to get $P(X_1, X_n = 0)$ we need $r(n-1)$ multiplications and $r(n-2)$ additions.

Additive decomposition

- The problem of finding the additive decomposition with minimal r corresponds to the problem of determining tensor rank.
- Determining tensor rank is an NP-hard problem.
- However, we have constructed explicit decomposition for some usefull probability distributions: noisy-or, noisy-and, noisy-add, noisy-max, noisy-min, etc.
- The above decompositions require low rank r, e.g., $r = 2$ for noisy-or and noisy-and,
- consequently, for these decompositions, computations of $P(X_i|X_j = x_j)$ is efficient - it has linear complexity with respect to n.

Additive decomposition

- The problem of finding the additive decomposition with minimal r corresponds to the problem of determining tensor rank.
- Determining tensor rank is an NP-hard problem.
- However, we have constructed explicit decomposition for some usefull probability distributions: noisy-or, noisy-and, noisy-add, noisy-max, noisy-min, etc.
- The above decompositions require low rank r, e.g., $r = 2$ for noisy-or and noisy-and,
- consequently, for these decompositions, computations of $P(X_i | X_j = x_j)$ is efficient - it has linear complexity with respect to n.

Additive decomposition

- The problem of finding the additive decomposition with minimal r corresponds to the problem of determining tensor rank.
- Determining tensor rank is an NP-hard problem.
- However, we have constructed explicit decomposition for some usefull probability distributions: noisy-or, noisy-and, noisy-add, noisy-max, noisy-min, etc.
- The above decompositions require low rank r, e.g., $r = 2$ for noisy-or and noisy-and,
- consequently, for these decompositions, computations of $P(X_i|X_j = x_j)$ is efficient - it has linear complexity with respect to n.

Additive decomposition

- The problem of finding the additive decomposition with minimal r corresponds to the problem of determining tensor rank.
- Determining tensor rank is an NP-hard problem.
- However, we have constructed explicit decomposition for some usefull probability distributions: noisy-or, noisy-and, noisy-add, noisy-max, noisy-min, etc.
- The above decompositions require low rank r, e.g., $r = 2$ for noisy-or and noisy-and,
- consequently, for these decompositions, computations of $P(X_i | X_j = x_j)$ is efficient - it has linear complexity with respect to n.

Additive decomposition

- The problem of finding the additive decomposition with minimal r corresponds to the problem of determining tensor rank.
- Determining tensor rank is an NP-hard problem.
- However, we have constructed explicit decomposition for some usefull probability distributions: noisy-or, noisy-and, noisy-add, noisy-max, noisy-min, etc.
- The above decompositions require low rank r, e.g., $r = 2$ for noisy-or and noisy-and,
- consequently, for these decompositions, computations of $P(X_i|X_j = x_j)$ is efficient - it has linear complexity with respect to n.