Markov and decomposability properties of the Kikuchi approximation

Roberto Santana, Pedro Larrañaga, José A. Lozano

Intelligent Systems Group

Department of Computer Science and Artificial Intelligence

University of the Basque Country

http://www.sc.ehu.es/isg/

Outline of the presentation

- 1. Region-based decompositions
- 2. Clique-based decompositions
- 3. Kikuchi approximation of the probability
- 4. Properties of the Kikuchi approximation
  - 4.1 Markov properties
  - 4.2 Decomposability properties
- 5. Partial Kikuchi approximations
- 6. Applications and future work

1. Region-based decompositions

- Region-based approximations of the energy
- Bethe approximation of the energy (Bethe, 1935)
- Kikuchi approximation of the energy (Kikuchi, 1951)
- Use of region-based decompositions in probabilistic inference (Yedidia, 2004)

1. Region-based decompositions

- Region R of the independence graph $G = (V, E)$
- Graph region-based decomposition $(\mathcal{R}, U)$
- Valid decomposition
- Clique-based decomposition

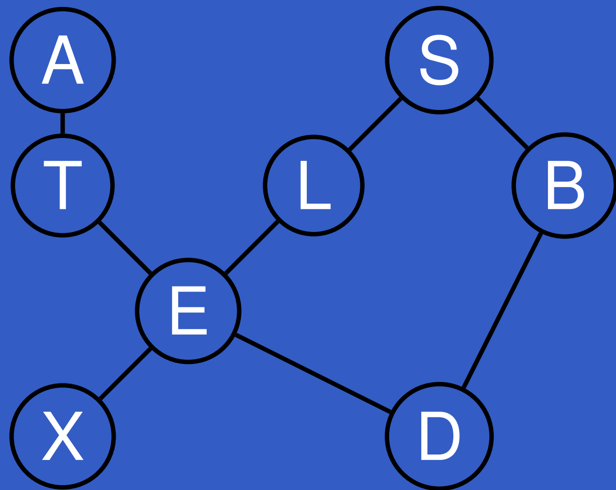
3. Kikuchi approximation

Kikuchi approximation of a probability distribution $p(\mathbf{x})$:

$$k(\mathbf{x}) = \prod_{R \in \mathcal{R}} p(\mathbf{x}_R)^{c_R}, \tag{1}$$

$$c_R = 1 - \sum_{\substack{S \in \mathcal{R} \\ R \subset S}} c_S \tag{2}$$

3. Kikuchi approximation



$$k(\mathbf{x}) = \frac{p(x_A, x_T)p(x_T, x_E)p(x_E, x_X)p(x_E, x_L)p(x_L, x_S)p(x_S, x_B)p(x_B, x_D)p(x_E, x_D)}{p(x_T)p(x_E)^3 p(x_L)p(x_S)p(x_B)p(x_D)} \tag{3}$$

4.1. Markov properties of the Kikuchi

- Pair-wise Markov property of the Kikuchi approximation is satisfied.
- Local Markov property of the Kikuchi approximation is satisfied.
- Global Markov property of the Kikuchi approximation is not satisfied.

4.1. Markov properties

Theorem 1 (Pairwise Markov property). *Given a Kikuchi approximation* $k(\mathbf{x})$ *defined on a graph* G*, and two variables* X_i *and* X_j*, if the corresponding vertices are not joined in* G*:*

$$k(x_i, x_j \mid \mathbf{x} \backslash (x_i, x_j)) = k(x_i \mid \mathbf{x} \backslash (x_i, x_j)) k(x_j \mid \mathbf{x} \backslash (x_i, x_j$$

4.1. Markov properties

Theorem 2 (Local Markov property). *Given a Kikuchi approximation $k(\mathbf{x})$ defined on a graph G, and a variable X_i,*

$$k(x_i, \mathbf{x} \backslash cl(x_i) \mid bd(x_i)) = k(x_i | bd(x_i)) k(\mathbf{x} \backslash cl(x_i) | bd(x_i))$$

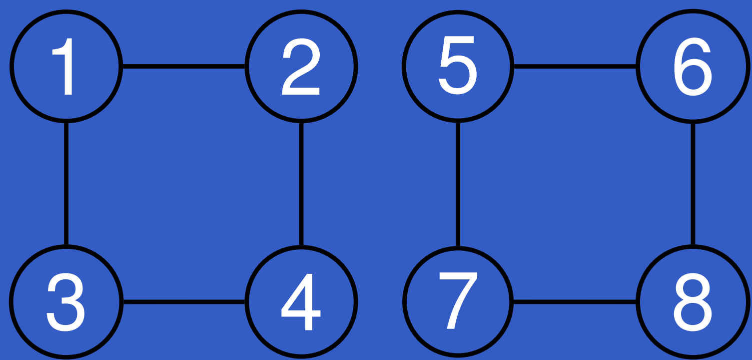
4.1. Markov properties

Conjecture 1 (Global Markov property). *Given a Kikuchi approximation* $k(\mathbf{x})$ *defined on a graph* G*, for all disjoint subsets* $\mathbf{X}_A$*,* $\mathbf{X}_B$*, and* $\mathbf{X}_C$*, whenever* $\mathbf{X}_B$ *and* $\mathbf{X}_C$ *are separated by* $\mathbf{X}_A$ *in the graph, then:*

$$k(\mathbf{x}_B, \mathbf{x}_C \mid \mathbf{x}_A) = k(\mathbf{x}_B \mid \mathbf{x}_A) k(\mathbf{x}_C \mid \mathbf{x}_A)$$

4.1. Markov properties.

Counterexample of conjecture 1. It corresponds a simple case, when $\mathbf{X}_B = X_i$, $\mathbf{X}_C = X_j$, X_i and X_j are not connected in the graph, and $\mathbf{X}_A = \emptyset$.



4.2. Decomposability properties

Theorem 3 (Kikuchi decomposition property). *Given a Kikuchi approximation* $k(\mathbf{x})$ *defined on a graph* G, *such that* $\mathbf{X} = \mathbf{X}_A \cup \mathbf{X}_B \cup \mathbf{X}_C$, *and* $\mathbf{X}_A$ *is a separator of* $\mathbf{X}_B$ *and* $\mathbf{X}_C$, *then:*

$$k(\mathbf{x}) = \frac{k_{AB}(\mathbf{x}_A, \mathbf{x}_B) k_{AC}(\mathbf{x}_A, \mathbf{x}_C)}{k_A(\mathbf{x}_A)} \tag{4}$$

4.2. Decomposability properties

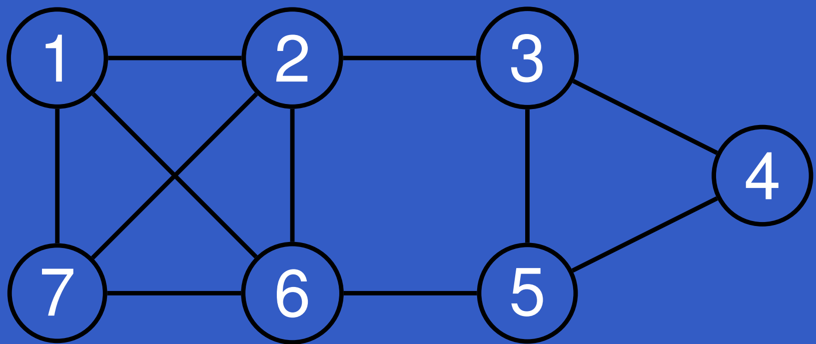
Definition 1. *There exists a decomposition of the random vector* $\mathbf{X}$ *with respect to a probability distribution* $p(\mathbf{x})$ *or equivalently,* $\mathbf{X}$ *is reducible, if and only if there exists a partition* $\mathbf{X}$ *into* $(\mathbf{X}_A, \mathbf{X}_B, \mathbf{X}_C)$ *such that:*

1. $p(\mathbf{x}_B, \mathbf{x}_C|\mathbf{x}_A) = p(\mathbf{x}_B|\mathbf{x}_A)p(\mathbf{x}_C|\mathbf{x}_A)$ *and neither* B *nor* C *are empty; and*
2. *the subgraph on* A*, in the independence graph of* $\mathbf{X}$ *is complete.*

If so, the components of $\mathbf{X}$ *are* $\mathbf{X}_{AB} = (\mathbf{X}_A, \mathbf{X}_B)$ *and* $\mathbf{X}_{AC} = (\mathbf{X}_A, \mathbf{X}_C)$*. If such a decomposition does not exist* $\mathbf{X}$ *is said to be irreducible.*

4.2. Decomposability properties

Example 1. *Consider the partition defined by sets* $A = \{2, 6\}$, $B = \{1, 7\}$ *and* $C = \{3, 4, 5\}$.



4.2. Decomposability properties of the

Theorem 4. *Given the independence graph G of $\mathbf{X}$, and the Kikuchi approximation $k(\mathbf{x})$ defined on G, if there exists a partition $\mathbf{X}$ into $(\mathbf{X}_A, \mathbf{X}_B, \mathbf{X}_C)$ such that the components of $\mathbf{X}$ are $\mathbf{X}_{AB} = (\mathbf{X}_A, \mathbf{X}_B)$ and $\mathbf{X}_{AC} = (\mathbf{X}_A, \mathbf{X}_C)$, then:*

$$k(\mathbf{x}) = \frac{k_{AB}(\mathbf{x}_{AB}) k_{AC}(\mathbf{x}_{AC})}{k_A(\mathbf{x}_A)} \quad (5)$$

5. Partial Kikuchi approximations

- Some of the irreducible incomplete components are approximated by the corresponding Kikuchi approximation.
- The rest of components are calculated exactly.

6. Applications and future work

- Definition of accuracy measures for Kikuchi approximations.
- Methods for learning Kikuchi approximations from data.
- Use of Kikuchi approximations in classification.
- Use of Kikuchi approximations in optimization.