Graph zeta function in the Bethe free energy

Physics of Algorithms

Santa Fe 2009/09/04

Yusuke Watanabe

(The Institute of Statistical Mathematics)

Joint work with my supervisor **Kenji Fukumizu**

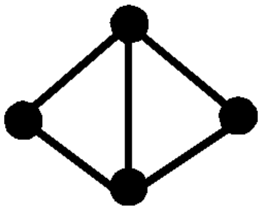
(The Institute of Statistical Mathematics)

outline

1. Intro
   - LBP and BFE
   - Graph Zeta function
2. The main formula
3. Implications

Intro: loopy belief propagation algorithm

- Pairwise Binary Model



$$G := (V, E) \qquad x_i \in \{\pm 1\}$$

$$p(x) = \frac{1}{Z} \prod_{ji \in E} \psi_{ji}(x_j, x_i)$$

- Loopy belief propagation algorithm (LBP)

$$m_{i \to j}(x_j) \propto \sum_{x_i = \pm} \psi_{ji}(x_j, x_i) \prod_{k \in N(i) \setminus \{j\}} m_{k \to i}(x_i)$$

$$\begin{cases} b_{ij}(x_i, x_j) \propto \psi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{k \to i}(x_i) \prod_{k \in N(j) \setminus i} m_{k \to j}(x_j), \\ b_i(x_i) \propto \prod_{k \in N(i)} m_{k \to i}(x_i), \end{cases}$$

Intro: Bethe free energy functional

- Definition of the Bethe free energy functional

$$F(b) := -\sum_{ij\in E}\sum_{x_i x_j} b_{ij}(x_i,x_j)\log\psi_{ij}(x_i,x_j) + \sum_{ij\in E}\sum_{x_i x_j} b_{ij}(x_i,x_j)\log b_{ij}(x_i,x_j) + \sum_{i\in V}(1-d_i)\sum_{x_i} b_i(x_i)\log b_i(x_i).$$

$d_i = |N(i)|$

$\{b_{ij}(x_i,x_j), b_i(x_i)\}$ Pseudo-marginals

Constraints: $b_{ij}(x_i,x_j) > 0$

$\sum_{x_j} b_{ij}(x_i,x_j) = b_i(x_i)$:Local consistency condition

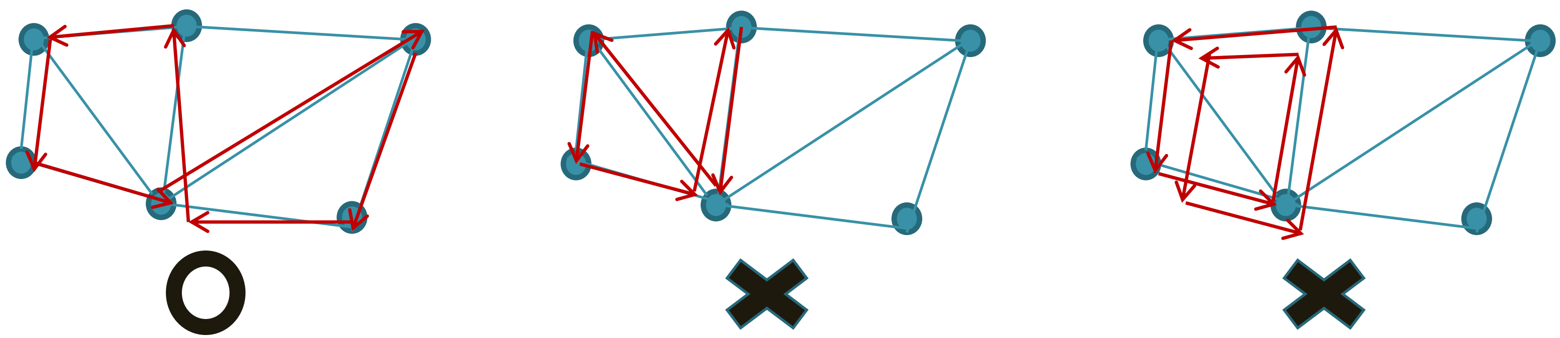
$\sum_{x_i,x_j} b_{ij}(x_i,x_j) = 1$:Normalization condition

- An important property of the Bethe free energy functional is that the set of **stationary points** corresponds to the set of **LBP fixed points**. (Yedidia et al, 2000)

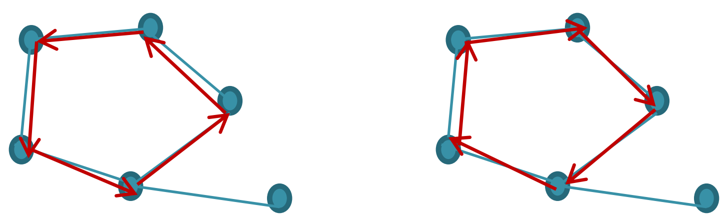
Intro: Graph zeta function (1/2)

- Definition of prime cycles .
  - A ***Prime cycle*** is a directed closed path that has **no back-tracking** and is **not a multiple** of shorter directed closed path.

    (We do not care about the starting point.)



  - <u>Example1</u>: For a **tree**, no prime cycles.
  - <u>Example2</u>: For a **1-cycle graph**, two prime cycles.



Intro: Graph zeta function (2/2)

- Definition of the (multivariable) ***graph zeta function***

  Assume that complex numbers u_e are associated with both directions of edges.

$$\zeta_G(\boldsymbol{u}) := \prod_{C \in P} (1 - g(C))^{-1}, \qquad \boldsymbol{u} = \{u_e\}$$

P : the set of prime cycles.

C : a prime cycle going through $e_1, \ldots, e_k$.

$g(C) := u_{e_1} \cdots u_{e_k}$

- Example1: For a **tree** $\zeta_G(\boldsymbol{u}) = 1$

- Example2: For a **1-cycle graph** $\zeta_G(\boldsymbol{u}) = (1 - \prod_{l=1}^{N} u_{e_l})^{-1}(1 - \prod_{l=1}^{N} u_{\bar{e}_l})^{-1}.$

  (The two prime cycles are $C = (e_1, \ldots, e_N)$ and $C' = (\bar{e}_N, \ldots, \bar{e}_1)$.

outline

1. Intro
   - LBP and BFE
   - Graph Zeta function
2. The main formula
3. Implications

The main formula

We use a parameterization of pseudomarginals:

$$b_{ij}(x_i, x_j) = \frac{1}{4}(1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j), \qquad b_i(x_i) = \frac{1}{2}(1 + m_i x_i)$$

The domain of the Bethe free energy functional is

$$L(G) := \left\{ \{m_i, \chi_{ij}\} \in \mathbb{R}^{|V|+|E|}; 1 + m_i x_i + m_j x_j + \chi_{ij} x_i x_j > 0 \text{ for all } ij \in E \text{ and } x_i, x_j = \pm 1 \right\}.$$

Theorem

At each point $\{m_i, \chi_{ij}\} \in L(G)$, the following formula holds.

$$\zeta_G(\boldsymbol{u})^{-1} = \det(\nabla^2 F) \prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j) \prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1 - d_i} \; 2^{2N + 4M}$$

where $\nabla^2 F$ is the Hessian of the BFE functional (w.r.t $\{m_i, \chi_{ij}\}$)
and $u_{i \to j} = \dfrac{\chi_{ij} - m_i m_j}{1 - m_j^2}$.

- **Remarks**
  - L.H.S. and R.H.S. are both functions of $\{m_i, \chi_{ij}\} \in L(G)$.
  - If the graph is a **tree**, L.H.S. equals to one.
  - This formula is generalized to more general models including the multinomial models and the Gaussian models.

outline

1. Intro
   - Loopy belief propagation algorithm
   - Bethe variational free energy
   - Graph Zeta function
2. The main formula
3. Implications
   - Local structure of LBP and BFE
   - Convexity condition
   - Other results from the formula

Implications 1: Local structures of LBP and BFE

Theorem

Let $\{b_{ij}(x_i, x_j), b_i(x_i)\}$ be a fixed point of LBP algorithm.
Let T be the linearization of the LBP update around the fixed point.
Then,

$$\det(I - T) = \det(\nabla^2 F) \prod_{ij \in E} \prod_{x_i, x_j = \pm 1} b_{ij}(x_i, x_j) \prod_{i \in V} \prod_{x_i = \pm 1} b_i(x_i)^{1-d_i} \; 2^{2N+4M}$$

- **Remarks**
  - This formula gives a concrete relation between local structure of the LBP update and the local structure of the BFE functional around a fixed point.
  - Locally stable fixed point of LBP $\Longrightarrow$ Local minima of the BFE functional (Multinomial models: proved by Heskes 2003. Zero-mean Gaussian models: our new result.)
  - For pairwise binary attractive models, we show that **"The temperature at which LBP becomes unstable coincides with the temperature at which the local minima of BFE functional disappears"**

Implications 2: Convexity condition

Lemma Let $\{m_i(t) := 0, \chi_{ij}(t) := t\} \in L(G)$ for $t < 1$

$$\lim_{t \to 1} \det(\nabla^2 F(t))(1-t)^{|E|+|N|-1} = -2^{-|E|-|V|+1}(|E| - |V|)\kappa(G),$$

where $\kappa(G)$ is the number of spanning trees in G.

• **Sketch of proof**

This formula is obtained by **Hashimoto's theorem**, which gives a limit of the (one variable) graph zeta function.

Theorem

The Bethe free energy functional is a convex function **if and only if** the underlying graph is a tree or 1-cycle graph.

• **Sketch of proof**

- "If" part is given by Pakzad, P., & Anantharam, V. (2002)
- "only if" part is obtained by the above lemma.

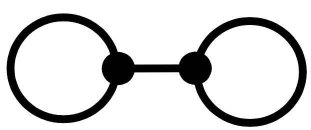
Implications 3: Other results

- The Hessian of the BFE functional **is positive definite** at $\{m_i, \chi_{ij}\} \in L(G)$

  if the **correlation coefficients** of the corresponding pseudomarginals are smaller than a certain value (which is determined by the graph topology).

- Consider a pairwise binary model on a connected graph with two linearly independent cycles.

  If the model is **not attractive** (under any gauge transformation) then the LBP fixed point is **unique**.



Thank you !