Young Scientist Meeting of Statistical Physics and Information Processing (YSM-SPIP), University of Tokyo, March 22-24, 2011

Connection between Free Energy and Belief Propagation on Random Factor Graph Ensembles

Ryuhei Mori

In this research, we deal with probability distributions defined by random sparse factor graphs. Free energy of random factor graph is informative in the context of low-density parity-check (LDPC) codes and constraint satisfaction problems (CSPs) [**1**]. Let G denote a factor graph consisting of N variable nodes and M factor nodes. Assume variable nodes take a value on a finite set $\mathcal{X}$. For each factor node a, there is a function $f_a(\boldsymbol{x}) : \mathcal{X}^{r_a} \to \mathbb{R}_{\geq 0}$ where r_a denotes degree of a factor node a. A probability distribution on $\mathcal{X}^N$ defined by a factor graph G is

$$p(\boldsymbol{x} \mid G) = \frac{1}{Z(G)} \prod_a f_a(\boldsymbol{x}_{\partial a})$$

where $\boldsymbol{x}_{\partial a}$ denotes value of variable nodes connecting to a factor node a, and where

$$Z(G) := \sum_{\boldsymbol{x} \in \mathcal{X}^N} \prod_a f_a(\boldsymbol{x}_{\partial a})$$

is a constant for normalization, called partition function. The purpose of this research is derivation of $\mathbb{E}[\log Z]$ by using the replica method where $\mathbb{E}[\cdot]$ denotes the expectation with respect to probability measure on a factor graph G.

In this abstract, the results for regular random factor graph ensemble are described. The result can be easily generalized to irregular and Poisson factor graph ensembles [**2**]. Let l and r denotes degrees of variable and factor nodes. Connection of edges is chosen uniformly from all $(Nl)!$ connections. For simplicity, it is assumed that $f_a(\boldsymbol{x})$ does not depend on a factor node a and denoted by $f(\boldsymbol{x})$. The basic idea of the calculation is type classification of $\boldsymbol{x} \in \mathcal{X}^N$ [**3**]. Let v denote the type of variable nodes i.e., the number of variable nodes of value $x \in \mathcal{X}$ is $v(x)$. Let u denote the type of factor nodes i.e., the number of factor nodes connecting to r variable nodes of value $\boldsymbol{x} \in \mathcal{X}^r$ is $u(\boldsymbol{x})$. Let $N(v, u \mid G)$ denote the number of assignments with type v and u on factor graph G. We can consider the classification according to the type of $\boldsymbol{x} \in \mathcal{X}^N$ in the partition function, namely,

$$Z(G) = \sum_{\boldsymbol{x} \in \mathcal{X}^N} \prod_a f(\boldsymbol{x}) = \sum_{v,u} N(v, u \mid G) \prod_{\boldsymbol{x} \in \mathcal{X}^r} f(\boldsymbol{x})^{u(\boldsymbol{x})}.$$

The expected number of assignments with type v and u is

$$\mathbb{E}[N(v, u \mid G)] = \binom{N}{\{v(x)\}_{x \in \mathcal{X}}} \binom{\frac{l}{r}N}{\{u(\boldsymbol{x})\}_{\boldsymbol{x} \in \mathcal{X}^r}} \frac{\prod_{x \in \mathcal{X}} (v(x)l)!}{(Nl)!}.$$

Now, we consider the exponent of the contribution of types ν and μ where $\nu(x) := v(x)/N$ and $\mu(x) := u(x)/((l/r)N)$, respectively. It holds

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}[Z(\nu, \mu)] = \frac{l}{r} \mathcal{H}(\mu) - (l-1)\mathcal{H}(\nu) + \frac{l}{r} \sum_{\boldsymbol{x} \in \mathcal{X}^r} \mu(\boldsymbol{x}) \log f(\boldsymbol{x}) =: -F_{\text{Bethe}}(\nu, \mu).$$

Note that F_{Bethe} has the similar form of Bethe free energy. It holds

$$\lim_{N \to \infty} \frac{1}{N} \log \mathbb{E}[Z(G)] = \max_{\nu, \mu} \{-F_{\text{Bethe}}(\nu, \mu)\}$$

where, ν and μ have to satisfy the following conditions.

$$\nu(x) \geq 0, \forall x \in \mathcal{X}, \qquad \mu(\boldsymbol{x}) \geq 0, \forall \boldsymbol{x} \in \mathcal{X}^r$$
$$\sum_{x\in\mathcal{X}} \nu(x) = 1, \qquad \sum_{\boldsymbol{x}\in\mathcal{X}^r} \mu(\boldsymbol{x}) = 1,$$
$$\frac{1}{r}\sum_{i=1}^{r}\sum_{\substack{\boldsymbol{x}\setminus x_i\\ x_i=z}} \mu(\boldsymbol{x}) = \nu(z), \forall z \in \mathcal{X}.$$

The last condition is for the consistency between ν and μ. The above maximization problem is similar to the minimization problem of Bethe free energy in which the stationary condition of Lagrangian is equivalent to the fixed point equation of belief propagation (BP) [**4**]. In the same way, the exponent of the moment $\mathbb{E}[Z(G)^n]$ can be calculated for $n \in \mathbb{N}$ since $Z(G)^n$ can be regarded as partition function of factor graph on alphabet $\mathcal{X}^n$ and factor $\prod_{i=1}^n f(\boldsymbol{x}^{(i)})$. Here, $\boldsymbol{x}^{(i)} \in \mathcal{X}^r$ denotes vector $(\boldsymbol{x}_1^{(i)}, \dots, \boldsymbol{x}_r^{(i)})$ where $\boldsymbol{x}_j$ is j-th elements of $\boldsymbol{x} \in (\mathcal{X}^n)^r$ and $\boldsymbol{x}_j^{(i)}$ denotes i-th element of $\boldsymbol{x}_j \in \mathcal{X}^n$.

THEOREM 1.
$$\lim_{N\to\infty} \frac{1}{N}\log \mathbb{E}[Z(G)^n] = \max_{(m_{\nu\to f}(\boldsymbol{x}), m_{f\to\nu}(\boldsymbol{x}))\in\mathcal{S}} \left\{ \frac{l}{r}\log Z_f + \log Z_\nu - l \log Z_{f\nu} \right\}.$$

where $\mathcal{S}$ denotes the set of saddle points of the function in max*, and where*

$$Z_\nu := \sum_{\boldsymbol{x}\in\mathcal{X}^n} m_{f\to\nu}(\boldsymbol{x})^l$$
$$Z_f := \sum_{\boldsymbol{x}\in(\mathcal{X}^n)^r} \left(\prod_{i=1}^n f(\boldsymbol{x}^{(i)})\right) \prod_{i=1}^r m_{\nu\to f}(\boldsymbol{x}_i)$$
$$Z_{f\nu} := \sum_{\boldsymbol{x}\in\mathcal{X}^n} m_{f\to\nu}(\boldsymbol{x}) m_{\nu\to f}(\boldsymbol{x}).$$

The conditions of saddle point are

$$m_{\nu\to f}(\boldsymbol{x}) \propto m_{f\to\nu}(\boldsymbol{x})^{l-1}$$
$$m_{f\to\nu}(\boldsymbol{x}) \propto \sum_{i=1}^r \sum_{\substack{\boldsymbol{x}\setminus\boldsymbol{x}_i\\ \boldsymbol{x}_i=\boldsymbol{x}}} f(\boldsymbol{x}) \prod_{j=1, j\neq i}^r m_{\nu\to f}(\boldsymbol{x}_j).$$

From this derivation, we can easily understand why BP equation appears in the calculation of exponent of moments since the problem is formulated as analogy to the minimization of Bethe free energy.

Replica symmetric assumption says that solutions $m_{f\to\nu}(x_1, \dots, x_n)$ and $m_{\nu\to f}(x_1, \dots, x_n)$ of the maximization problem are invariant under permutations on x_1 to x_n. Furthermore, the representations

$$m_{\nu\to f}(\boldsymbol{x}) = \int \prod_{i=1}^n M_{\nu\to f}(x_i) \mathrm{d}\Phi(M_{\nu\to f})$$
$$m_{f\to\nu}(\boldsymbol{x}) = \int \prod_{i=1}^n M_{f\to\nu}(x_i) \mathrm{d}\hat{\Phi}(M_{f\to\nu})$$

are assumed where Φ and $\hat{\Phi}$ denote probability measures on $\mathcal{P}(\mathcal{X})$, i.e., Φ and $\hat{\Phi}$ are elements of $\mathcal{P}(\mathcal{P}(\mathcal{X}))$. Here, $\mathcal{P}(\mathcal{A})$ denotes the set of probability measures on a set $\mathcal{A}$.

LEMMA 2.
$$-F_{RS} = \max_{(\Phi,\hat{\Phi})\in\mathcal{S}} \left\{ \frac{l}{r}\langle \log \mathcal{Z}_f\rangle + \langle \log \mathcal{Z}_\nu\rangle - l\langle \log \mathcal{Z}_{f\nu}\rangle \right\}$$

where $\mathcal{S}$ denotes the set of saddle points of the function in max*, where*

$$\mathcal{Z}_v := \sum_{x\in\mathcal{X}} \prod_{i=1}^{l} M^{(i)}_{f\to v}(x)$$

$$\mathcal{Z}_f := \sum_{\boldsymbol{x}\in\mathcal{X}^r} f(\boldsymbol{x}) \prod_{i=1}^{r} M^{(i)}_{v\to f}(x_i)$$

$$\mathcal{Z}_{fv} := \sum_{x\in\mathcal{X}} M_{v\to f}(x) M_{f\to v}(x)$$

where $\{M^{(i)}_{v\to f}\}_{i=1,\cdots,r}$ and $\{M^{(i)}_{f\to v}\}_{i=1,\cdots,l}$ are i.i.d. random measures obeying Φ and $\hat{\Phi}$, respectively, and where $\langle\cdot\rangle$ denotes the expectation with respect to the random measures. The saddle point conditions are

$$\frac{\prod_{i=1}^{l-1} M^{(i)}_{f\to v}(x)}{\sum_{x\in\mathcal{X}} \prod_{i=1}^{l-1} M^{(i)}_{f\to v}(x)} \sim \Phi$$

$$\frac{\sum_{\boldsymbol{x}\in\mathcal{X}^r, x_D = x} f(\boldsymbol{x}) \prod_{j=1, j\neq D}^{r} M^{(j)}_{v\to f}(x_j)}{\sum_{\boldsymbol{x}\in\mathcal{X}^r} f(\boldsymbol{x}) \prod_{j=1, j\neq D}^{r} M^{(j)}_{v\to f}(x_j)} \sim \hat{\Phi}$$

where D denotes the uniform random variable on $\{1,2,\dots,r\}$ which is independent of any random variable, and where $M \sim \Phi$ denotes that a random measure M has a law Φ.

This derivation of replica symmetric solution is simpler than previously known ones in which complicated tools are used [**5**] e.g., integral expression of the delta function. Another advantage of this research is that we can understand why the saddle point equation in the replica symmetric solution is equal to the DE equation.

When $f(\boldsymbol{x})$ is invariant under permutation on $\boldsymbol{x}$, the fixed points for annealed free energy in Theorem 1 for $n = 1$ are also fixed point for RS saddle point equation as delta distribution. From the inclusion relation of domains of max in Theorem 1 and Lemma 2, $-F_{\rm RS} \geq \lim_{N\to\infty} 1/N \log \mathbb{E}[Z]$. On the other hand, from Jensen's inequality, $\mathbb{E}[\log Z] \leq \log \mathbb{E}[Z]$. We now obtain the following theorem.

THEOREM 3. *Assume $f(\boldsymbol{x})$ is invariant under permutation on $\boldsymbol{x}$. If replica symmetric assumption is valid i.e., $-F_{RS} = \lim_{N\to\infty} 1/N\mathbb{E}[\log Z]$, then $\lim_{N\to\infty} 1/N\mathbb{E}[\log Z] = \lim_{N\to\infty} 1/N \log \mathbb{E}[Z]$.*

This result is well known for regular LDPC codes [**5**].

Acknowledgment

The author acknowledges Toshiyuki Tanaka for insightful discussion. This work was supported by the Grant-in-Aid for Scientific Research for JSPS Fellows (22·5936), MEXT, Japan.

References

[1] M. Mezard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, USA, 2009.

[2] R. Mori, "Connection between annealed free energy and belief propagation on random factor graph ensembles," 2011. [Online]. Available: http://arxiv.org/abs/1102.3132v2

[3] P. Vontobel, "Counting in graph covers: A combinatorial characterization of the bethe entropy function," 2010. [Online]. Available: http://arxiv.org/abs/1012.0065

[4] J. Yedidia, W. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.

[5] S. Condamin, "Study of the weight enumerator function for a Gallager code," *Project Report Cavendish Laboratory, University of Cambridge, Cambridge, UK*, 2002.

SYSTEMS SCIENCE, DEPARTMENT OF INFORMATICS, KYOTO UNIVERSITY
E-mail address, R. Mori: `rmori@sys.i.kyoto-u.ac.jp`