Free energy

February 12, 2021

Belief propagation equations:

$$\begin{aligned}
b_i(\sigma_i) &= \frac{1}{z_i}\prod_{a\in\partial i} m_{ai}(\sigma_i) \\
b_a(\underline{\sigma}_a) &= \frac{1}{z_a}\Psi_a(\underline{\sigma}_a)\prod_{i\in\partial a} m_{ia}(\sigma_i) \\
m_{ai}(\sigma_i) &= \frac{1}{z_{a\to i}}\sum_{\underline{\sigma}_{a\setminus i}}\Psi_a(\underline{\sigma}_a)\prod_{j\in\partial a\setminus i} m_{ja}(\sigma_i) \\
m_{ia}(\sigma_i) &= \frac{1}{z_{i\to a}}\prod_{b\in\partial i\setminus a} m_{bi}(\sigma_i)
\end{aligned} \tag{1}$$

1 Re-define message normalizations

Define

$$z_{ai} \coloneqq \sum_{\sigma_i} m_{ai}(\sigma_i)\, m_{ia}(\sigma_i) \tag{2}$$

So that now

$$\begin{aligned}
b_i(\sigma_i) &= \sum_{\underline{\sigma}_{a\setminus i}} b_a(\underline{\sigma}_a) \\
&= \sum_{\underline{\sigma}_{a\setminus i}}\frac{1}{z_a}\Psi_a(\underline{\sigma}_a)\prod_{i\in\partial a} m_{ia}(\sigma_i) \\
&= \frac{1}{z_a}\underbrace{\sum_{\underline{\sigma}_{a\setminus i}}\Psi_a(\underline{\sigma}_a)\prod_{j\in\partial a\setminus i} m_{ja}(\sigma_j)}_{z_{a\to i}m_{ai}(\sigma_i)}\, m_{ia}(\sigma_i) \\
&= \frac{z_{a\to i}}{z_a} m_{ai}(\sigma_i)\, m_{ia}(\sigma_i)
\end{aligned} \tag{3}$$

Since $b_i(\sigma_i)$ is normalized by construction,

$$1 = \sum_{\sigma_i} b_i(\sigma_i) = \frac{z_{a\to i}}{z_a} z_{ai} \tag{4}$$

and therefore

$$\boxed{z_{a\to i} = \frac{z_a}{z_{ai}}} \tag{5}$$

Similarly, in order to express also $z_{i\to a}$ in terms of z_{ai}, one can re-write the expression for b_i separating the a-th incoming message

$$
\begin{aligned}
b_i\left(\sigma_i\right) &= \frac{1}{z_i} \underbrace{\prod_{b\in\partial i\backslash a} m_{bi}\left(\sigma_i\right)}_{z_{i\to a} m_{ia}\left(\sigma_i\right)} m_{ai}\left(\sigma_i\right) \\
&= \frac{z_{i\to a}}{z_i} m_{ia}\left(\sigma_i\right) m_{ai}\left(\sigma_i\right)
\end{aligned} \tag{6}
$$

Again imposing normalization for b_i,

$$
1 = \sum_{\sigma_i} b_i\left(\sigma_i\right) = \frac{z_{i\to a}}{z_i} z_{ai} \tag{7}
$$

giving

$$
\boxed{z_{i\to a} = \frac{z_i}{z_{ai}}} \tag{8}
$$

The message updates become

$$
\begin{aligned}
m_{ai}\left(\sigma_i\right) &= \frac{z_{ai}}{z_a} \sum_{\underline{\sigma}_{a\backslash i}} \Psi_a\left(\underline{\sigma}_a\right) \prod_{j\in\partial a\backslash i} m_{ja}\left(\sigma_i\right) \\
m_{ia}\left(\sigma_i\right) &= \frac{z_{ai}}{z_i} \prod_{b\in\partial i\backslash a} m_{bi}\left(\sigma_i\right)
\end{aligned} \tag{9}
$$

2 Bethe Free energy in terms of messages

The Bethe free entropy is given by

$$
-\beta F(\underline{b}) = \sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \log\left[\frac{b_a\left(\underline{\sigma}_a\right)}{\Psi_a\left(\underline{\sigma}_a\right)}\right] + \sum_i \left(1 - |\partial i|\right) \sum_{\sigma_i} b_i\left(\sigma_i\right) \log\left[b_i\left(\sigma_i\right)\right] \tag{10}
$$

where F is the free energy. (We compute the free entropy for easier comparison with [1]).

We would like instead an expression $F(\underline{m})$ only in terms of the messages. This is found to be a more convenient form when one wants to, for instance, compute the average free energy over distributions of messages obtained via the cavity method.

Substituting the definitions for node and factor beliefs in the logs,

$$
\begin{aligned}
\log\left[\frac{b_a\left(\underline{\sigma}_a\right)}{\Psi_a\left(\underline{\sigma}_a\right)}\right] &= \log\left[\frac{1}{z_a} \prod_{i\in\partial a} m_{ia}\left(\sigma_i\right)\right] \\
&= \sum_{i\in\partial a} m_{ia}\left(\sigma_i\right) - \log z_a
\end{aligned} \tag{11}
$$

$$
\log b_i\left(\sigma_i\right) = \sum_{a\in\partial i} \log\left(m_{ai}\left(\sigma_i\right)\right) - \log z_i \tag{12}
$$

The first term in (10) is

$$\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \log\left[\frac{b_a\left(\underline{\sigma}_a\right)}{\Psi_a\left(\underline{\sigma}_a\right)}\right] = \sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \log m_{ia}\left(\sigma_i\right) + \sum_a \underbrace{\sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right)}_{1} \log z_a$$
$$= \sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \log m_{ia}\left(\sigma_i\right) + \sum_a \log z_a \tag{13}$$

The second is

$$\sum_i \left(1-|\partial i|\right) \sum_{\sigma_i} b_i\left(\sigma_i\right) \log\left[b_i\left(\sigma_i\right)\right] =$$
$$\sum_i \left(1-|\partial i|\right) \left[\sum_{\sigma_i} b_i\left(\sigma_i\right) \sum_{a\in\partial i} \log m_{ai}\left(\sigma_i\right) + \underbrace{\sum_{\sigma_i} b_i\left(\sigma_i\right)}_{1} \log z_i\right] =$$
$$\sum_i \left(1-|\partial i|\right) \left[\sum_{\sigma_i} b_i\left(\sigma_i\right) \sum_{a\in\partial i} \log m_{ai}\left(\sigma_i\right) + \log z_i\right] \tag{14}$$

The free entropy is now

$$-\beta F = \sum_a \log z_a + \sum_i \left(1-|\partial i|\right) \log z_i -$$
$$\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \log m_{ia}\left(\sigma_i\right) - \sum_i \left(1-|\partial i|\right) \sum_{\sigma_i} b_i\left(\sigma_i\right) \sum_{a\in\partial i} \log m_{ai}\left(\sigma_i\right) \tag{15}$$

Re-write

$$\sum_i |\partial i| \log z_i = \sum_i |\partial i| \log\left[\sum_{\sigma_i} \prod_{b\in\partial i} m_{bi}\left(\sigma_i\right)\right]$$
$$= \sum_i \sum_{a\in\partial i} \log\left[\sum_{\sigma_i} m_{ai} \underbrace{\prod_{b\in\partial i\setminus a} m_{bi}\left(\sigma_i\right)}_{m_{ia}\left(\sigma_i\right) z_{i\to a}}\right]$$
$$= \sum_i \sum_{a\in\partial i} \log\left[\underbrace{\sum_{\sigma_i} m_{ai} m_{ia}\left(\sigma_i\right) z_{i\to a}}_{z_{ai}\ \text{(from (2))}}\right]$$
$$= \sum_{\langle i,a\rangle} \left(\log z_{ai} + \log z_{i\to a}\right) \tag{16}$$

Now

$$
\begin{aligned}
-\beta F = &\sum_a \log z_a + \sum_i \log z_i - \sum_{\langle i,a\rangle} \log z_{ai} - \sum_{\langle i,a\rangle} \log z_{i\to a} \\
&- \sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \log m_{ia}\left(\sigma_i\right) - \sum_i \left(1-|\partial i|\right) \sum_{\sigma_i} b_i\left(\sigma_i\right) \sum_{a\in\partial i} \log m_{ai}\left(\sigma_i\right)
\end{aligned}
\tag{17}
$$

Expanding $\log m_{ia}\left(\sigma_i\right)$ in the fourth term,

$$
\begin{aligned}
&\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \log m_{ia}\left(\sigma_i\right) = \\
&\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \log m_{bi}\left(\sigma_i\right) - \sum_a \underbrace{\sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right)}_{1} \sum_{i\in\partial a} \log z_{i\to a} = \\
&\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \log m_{bi}\left(\sigma_i\right) - \sum_{\langle i,a\rangle} \log z_{i\to a}
\end{aligned}
\tag{18}
$$

therefore

$$
\begin{aligned}
-\beta F = &\sum_a \log z_a + \sum_i \log z_i - \sum_{\langle i,a\rangle} \log z_{ai} - \cancel{\sum_{\langle i,a\rangle} \log z_{i\to a}} \\
&- \sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \log m_{bi}\left(\sigma_i\right) + \cancel{\sum_{\langle i,a\rangle} \log z_{i\to a}} \\
&- \sum_i \left(1-|\partial i|\right) \sum_{\sigma_i} b_i\left(\sigma_i\right) \sum_{a\in\partial i} \log m_{ai}\left(\sigma_i\right)
\end{aligned}
\tag{19}
$$

Now the goal is to get rid of the fourth and fifth term. We re-arrange the sums in the fourth term in order to recover the definition of b_i as marginalization of b_a

$$
\begin{aligned}
-\sum_a \sum_{\underline{\sigma}_a} b_a\left(\underline{\sigma}_a\right) \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \log m_{bi}\left(\sigma_i\right) &= -\sum_a \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \sum_{\sigma_i} \underbrace{\sum_{\underline{\sigma}_{a\setminus i}} b_a\left(\underline{\sigma}_a\right)}_{b_i\left(\sigma_i\right)} \log m_{bi}\left(\sigma_i\right) \\
&= -\sum_a \sum_{i\in\partial a} \sum_{b\in\partial i\setminus a} \sum_{\sigma_i} b_i\left(\sigma_i\right) \log m_{bi}\left(\sigma_i\right)
\end{aligned}
$$

We can substitute $\sum_a \sum_{i\in\partial a}$ with $\sum_i \sum_{a\in\partial i}$ since they both count each edge exactly once

$$
= -\sum_i \sum_{a\in\partial i} \sum_{b\in\partial i\setminus a} \sum_{\sigma_i} b_i\left(\sigma_i\right) \log m_{bi}\left(\sigma_i\right)
\tag{20}
$$

Now the fourth and fifth terms of (19) look like

$$
-\sum_i \sum_{a\in\partial i} \sum_{b\in\partial i\setminus a} \sum_{\sigma_i} b_i\left(\sigma_i\right) \log m_{bi}\left(\sigma_i\right) - \sum_i \left(1-|\partial i|\right) \sum_{a\in\partial i} \sum_{\sigma_i} b_i\left(\sigma_i\right) \log m_{ai}\left(\sigma_i\right)
\tag{21}
$$

which gives zero, because the quantity $\sum_{\sigma_i} b_i\left(\sigma_i\right) \log m_{bi}\left(\sigma_i\right)$ is counted $\left(|\partial i|-1\right)$ times by the summation $\sum_{b\in\partial i\setminus a}$ over all neighbors of i except one (a).

Finally, we get the free entropy in terms of messages

$$\boxed{-\beta F(\underline{m}) = \sum_a \log z_a(\underline{m}) + \sum_i \log z_i(\underline{m}) - \sum_{\langle i,a \rangle} \log z_{ai}(\underline{m})} \tag{22}$$

References

[1] M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, 2009.