Comparison BP, MF and exact on Ising model

In this exercise you will assess the accuracy of the mean field approximation and belief propagation on on the Ising model, by comparing with the exact results.

Ising model

The Ising model is defined as $p(x) = \frac{1}{Z} \exp\left(\sum_{(ij)} w_{ij} x_i x_j + \sum_i \theta_i x_i\right)$. The couplings w_{ij} are defined on a sparse graph. The weights are random from a Gaussian distribution with mean J_0/k and variance J^2/k with $k = cn$, n the number of spins and c the density of connections (there are approximately cn^2 non-zero connections in w. The θ_i are also random.

Mean field (MF) approximation

The mean field approximation $m_i^{\mathrm{MF}} = m_i$ is defined as the solution to the system of equations $m_i = \tanh\left(\sum_{j \in N(i)} w_{ij} m_j + \theta_i\right)$. You compute the solution using so-called fixed point iteration. Start with a random vector m, compute the rhs of the equation $m' = \tanh\left(\sum_{j \in N(i)} w_{ij} m_j + \theta_i\right)$ and set $m = m'$. In each iteration you check how much m changes as `dm=max(abs(m-m_old))`. Put the iterations in a while loop and stop when $dm < \epsilon$ with ϵ is of the order of the machine precision or 1×10^{-13}.

Convergence should occur in less than several hundred iterations provided that the w_{ij} are sufficiently small. For larger w_{ij} convergence slows down or might break down entirely. The fixed point iteration starts to oscillate and does not converge. When this arises, a good solution is called smoothing. Replace the fixed point iteration updates by

$$m := \eta m + (1 - \eta) \tanh\left(\sum_{j \in N(i)} w_{ij} m_j + \theta_i\right) \qquad 0 \leq \eta < 1$$

The new m is a linear combination of the suggested update $\tanh\left(\sum_{j \in N(i)} w_{ij} m_j + \theta_i\right)$ and the old m. Smoothing has better convergence properties but converges slower.

Belief propagation (BP)

For BP, we need to compute the solution of the message passing equation

$$m_{ij}(x_j) \propto \sum_{x_i} \psi_{ij}(x_i, x_j) \psi_i(x_i) \prod_{k \in N(i) \setminus j} m_{ki}(x_i)$$

Each message $m_{ij}(x_j)$ can be treated as a probability distribution over a single binary variable and is therefore fully defined by a single number. Convenient parametrizations are

$$m_{ij}(x_j) = \frac{1}{2}(1 + \mu_{ij} x_j) \propto e^{a_{ij} x_j}$$

$\mu_{ij} = \sum_{x_j} x_j m_{ij}(x_j)$ is the expected value and a_{ij} is the parameter if we write $m_{ij}(x_j)$ as a exponential distribution. It is easy to show that $\mu_{ij} = \tanh(a_{ij})$. In terms of a_{ij} the BP equation becomes

$$m_{ij}(x_j) = \sum_{x_i} \exp\left(w_{ij}x_i x_j + \theta_i x_i + \sum_{k\in N(i)\setminus j} a_{ki}x_i\right) \tag{1}$$

$$= 2\cosh\left(w_{ij}x_j + \theta_i + \sum_{k\in N(i)\setminus j} a_{ki}\right) \tag{2}$$

Note, that $m_{ij}(x_j)$ defined in this way is not yet normalized. Thus, we obtain [1]

$$a_{ij} = \tanh^{-1}(\mu_{ij}) = \tanh^{-1}\left(\frac{m_{ij}(x_j=1) - m_{ij}(x_j=-1)}{m_{ij}(x_j=1) + m_{ij}(x_j=-1)}\right) = \frac{1}{2}\log\frac{m_{ij}(x_j=1)}{m_{ij}(x_j=-1)} \tag{3}$$

Thus, the BP solution is computed as a fixed point iteration in the matrix a. That is to say, in each iteration, the entire matrix is updated. The pseudo code looks like

```
Start with a random initial n × n matrix a
da = 1
while da > 1 × 10^-13 do
    a_old = a
    compute matrix m_ij(x_j = 1) and m_ij(x_j = −1) according to Eq. 2
    compute matrix a according to Eq. 3
    da = max(max(abs(a − a_old)))
end while
m_i = tanh(θ_i + Σ_j a_ji)                    ▷ BP approximation for mean spin values
```

Exercises

1. You are required to write your own code in your preferred language. I have provided a Matlab template `main.m` that contains the definition of the Ising model and its parameters and the code for the exact computation. Write your code for the MF approximation and the BP approximation.

2. Consider the fully connected Ising model ($c = 1$) with $n = 20, J_0 = 0, J_{\text{th}} = J = \beta$.

   (a) Compute the RMS error in the mean values $\langle x_i \rangle$ as $R_{\text{MF}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(m_i^{\text{ex}} - m_i^{\text{mf}}\right)^2}$ with $m_i^{\text{ex}} = \langle x_i \rangle$ the exact spin expectation values and m_i^{MF} their MF approximations, and similar for BP. For each β, estimate the RMS errors in multiple random instances of the Ising model and compute the mean and standard deviation. Plot the RMS errors (mean and standard deviation) as a function of β for $0 < \beta \leq 2$. You should find that the algorithms converge in several hundred iterations using smoothing $\eta = 0.5$. You should find that the MF and BP results look somewhat like fig. 1.

   (b) Compute the RMS error in the correlations $\chi_{ij} = \left\langle x_i x_j \right\rangle - \langle x_i \rangle \left\langle x_j \right\rangle$ using the MF approximation with linear response correction χ_{ij}^{LR} and χ_{ij}^{BP} from BP. The latter is computed as follows. After convergence, the joint probability

[1] In the last step we use the identity $\frac{1}{2}\log\frac{1+m}{1-m} = \tanh^{-1}(m)$.

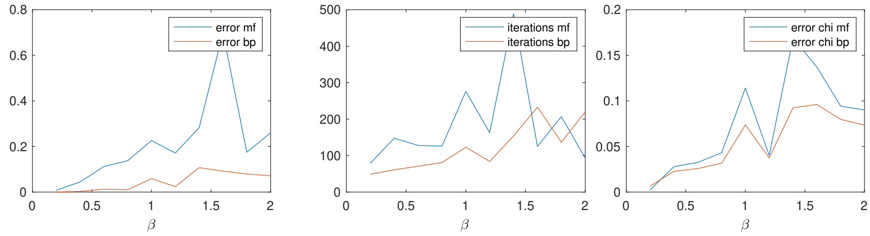

Figure 1: Accuracy of MF and BP approximation for fully connected model with $n = 20$ versus β. Left: RMS errors in mean m_i. Middle: number of iterations until convergence. Right: RMS error in connected correlations χ. Smoothing $\eta = 0.5$.

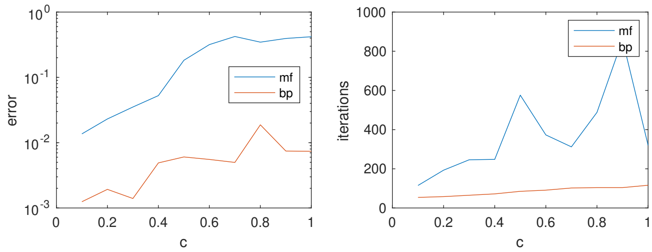

Figure 2: Errors and number of iterations for MF and BP on sparse model versus connectivity c. $n = 20, \theta = 0.1, \beta = 0.2$, smoothing for both MF and BP is $\eta = 0.5$

of two nodes x_i, x_j is

$$\begin{aligned} b_{ij}(x_i, x_j) &= \frac{1}{Z} \exp\left(w_{ij} x_i x_j + \theta_i x_i + \theta_j x_j + \sum_{k \in N(i) \setminus j} a_{ki} x_i + \sum_{l \in N(j) \setminus i} a_{lj} x_j \right) \\ Z &= \sum_{x_i, x_j} \exp\left(w_{ij} x_i x_j + \theta_i x_i + \theta_j x_j + \sum_{k \in N(i) \setminus j} a_{ki} x_i + \sum_{l \in N(j) \setminus i} a_{lj} x_j \right) \end{aligned}$$

and $\chi_{ij}^{BP} = \sum_{x_i, x_j} x_i x_j b_{ij}(x_i, x_j) - m_i^{\mathrm{BP}} m_j^{\mathrm{BP}}$. For each β, estimate the RMS errors between χ_{ij}^{LR} and the exact values χ_{ij}^{ex}, and between χ_{ij}^{BP} and the exact values χ_{ij}^{ex} in multiple random instances of the Ising model and compute the mean and standard deviation. Plot the RMS errors (mean and standard deviation) as a function of β for $0 < \beta \leq 2$. See fig. 1.

3. Consider the sparse Ising model with $n = 20$ and sparsity c with non-zero couplings $w = \pm\beta$. [2] Put all threshold values equal to $\theta > 0$. Plot the RMS errors (mean and standard deviation) as a function of c, β for $0 < c \leq 1$ and $0 < \beta \leq 1$ and for various θ. [3] You should find that the BP results are more accurate than MF for sparse networks. For instance see fig. 2

[2] I used the Matlab command `w=sprandsym(n,c)`. This should give a symmetric weight matrix w with $c * n^2$ non-zero elements. However, I find that the average fraction of non-zero elements in w tends to be lower. For instance $c = 1$ in fig. 2 actually corresponds to approximately 60 % non-zero elements in w.

[3] Do not present 3d plots (c, β, R) because these tend to be hard to interpret. It is usually clearer to present dependence on β for various c and dependence on c for various β.

References

[Kappen and Spanjers, 1999] Kappen, H. and Spanjers, J. (1999). Mean field theory for asymmetric neural networks. *Physical Review E*, 61:5658–5663.