Excercise Computational neuroscience: Boltzmann-Gibbs distributions

1. (a) Derive Eq. **??**.

   (b) Show that the detailed balance does not hold when the weights of the neural network are not symmetric ($w_{ij} \neq w_{ji}$). In other words, show that te Boltzmann distribution is not the stationary distribution of the Glauber dynamics with asymmetric weights.

2. Study the accuracy of the mean field and linear response method for a Boltzmann distribution on 2 neurons with equal threshold $\theta_1 = \theta_2 = \theta$ and connected by a weigth w:

$$p(s_1, s_2) = \frac{1}{Z} \exp(w s_1 s_2 + \theta(s_1 + s_2))$$

   (a) Give an expression for the mean field equations to approximately compute the firing rates for this network.

   (b) Solve the mean field equations numerically for $\theta = w$ and various values of w and compare the mean field approximation with the exact result.

   (c) Compute the linear response estimate of the correlations and compare with the exact values.

3. Work out analytically the result of mean field learning with linear response correction for the case of two neurons and a data set consisting of three patterns $(1, -1), (1, 1), (-1, -1)$.

1 Answers

1. (a) From the definition of the sequential Glauber dynamics:

$$T(s|F_i s) = \sigma(s_i h_i(F_i s)) = \sigma(s_i h_i(s))$$

The last step follows becasue $F_i s$ is the state s with the value of the ith neuron flipped ($s_i \rightarrow -s_i$). The local field h_i is given by

$$h_i = \sum_{j \neq i} w_{ij} s_j + \theta_i$$

and therefore does not depend on the value of s_i.
The reverse transition probability is

$$T(F_i s|s) = \sigma(-s_i h_i(s))$$

Therefore,

$$\begin{aligned} \frac{T(s|F_i s)}{T(F_i s|s)} &= \frac{\sigma(s_i h_i(s))}{\sigma(-s_i h_i(s))} \\ &= \frac{\exp(s_i h_i(s))}{\exp(h_i(s)) + \exp(-h_i(s))} \frac{\exp(h_i(s)) + \exp(-h_i(s))}{\exp(-s_i h_i(s))} \\ &= \exp(2 s_i h_i(s)) \end{aligned}$$

(b) The Glauber dynamics with sequential update satisfies Eq. **??** also for asymmetric weights. On the other hand, the Boltzmann distribution is independent of the asymmetric part of the weights. In other words, if we write $w_{ij} = w_{ij}^s + w_{ij}^a$ with $w_{ij}^{s,a}$ the symmetric and asymmetric parts of the matrix w_{ij}, respectively. We have

$$\sum_{ij} w_{ij} s_i s_j = \sum_{ij} w_{ij}^s s_i s_j$$

Therefore Eq. **??** and Eq. **??** can only be reconciled for symmetric w_{ij}.

2. (a) The mean field equations are given by

$$m_1 = \tanh(wm_2 + \theta), \quad m_2 = \tanh(wm_1 + \theta)$$

When $w > 0$, the solution is of the form $m = m_1 = m_2$ and therefore we only have to solve

$$m_{mf} = \tanh(wm_{mf} + \theta)$$

for m_{mf}.

(b) The solution of this fixed point equation for $\theta = w$ is sketched in fig. 1Left. For any value of w, we can solve this equation using any standard numerical routine. The Matlab code that I used is given here:

```
m=-2:0.1:2;
subplot(2,2,1)
plot(m,m,m,tanh(m+1),m,tanh(10*m+10))
legend('m','tanh(m+1)','tanh(10m+10)')
xlabel('m')
i=0;
for w=0:0.1:1,
i=i+1;
m=fsolve(inline('tanh(w*m+w)-m','m','w'),0,[],w);
w1(i)=w;
m_mf(i)=m;
m_ex(i)=(-exp(-w)+exp(3*w))/(3*exp(-w)+exp(3*w));
end;
subplot(2,2,2)
plot(w1,m_mf,w1,m_ex)
legend('m_{mf}','m_{ex}')
xlabel('w')
ylabel('m')
```

The exact mean firing rates are normally intractable to compute for a large network, but for two neurons is easily computed. It is given by Eq. **??**, which for the special case that $\theta = w$ becomes

$$m_{ex} = \frac{-\exp(-w) + \exp(3w)}{3\exp(-w) + \exp(3w)}$$

For $\theta = w$, the solution for various values of w is shown in fig.1Right.

(c) The exact correlations are given by Eq. **??**, with

$$\langle s_1 s_2 \rangle = \sum_{s_1, s_2} s_1 s_2 p(s) = \frac{-\exp(-w) + \exp(3w)}{3\exp(-w) + \exp(3w)}$$

and $\langle s_i \rangle = m_{ex}$ as above. The linear response estimate of the correlations is given by Eq. **??**, with m_i the mean field solution. Thus, for each w and θ, we first compute the mean field solution m, and then invert the matrix in Eq. **??**. The exact and linear response solution for $\theta = w$ as a function of w is plotted in fig. 2. The Matlab code that I used is given here:

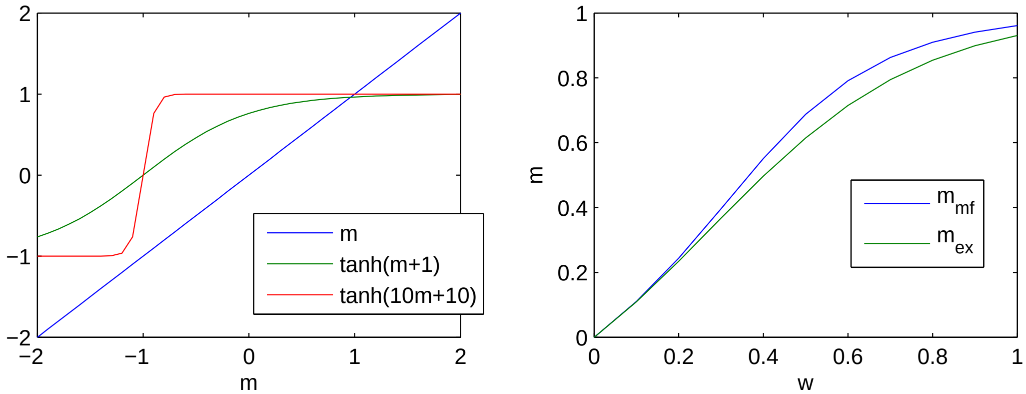

Figure 1: Left) Solution of the fixed point equation is given at the intersection of the straight line and the tanh. Right) Numerical solutions are obtained solving the one dimensional non-linear equations.

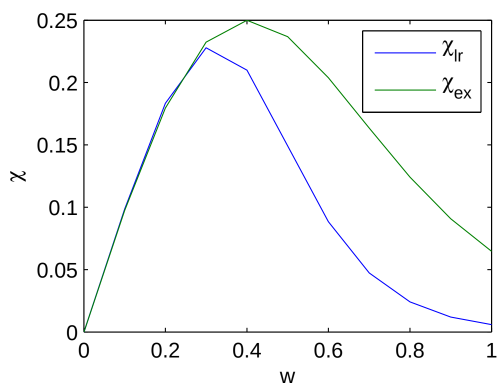

Figure 2: Comparison of linear response and exact correlations.

```
i=0;
for w=0:0.1:1,
    i=i+1;
    w1(i)=w;
    m=fsolve(inline('tanh(w*m+w)-m','m','w'),0,[],w);
    c=[1/(1-m^2), -w;-w,1/(1-m^2)];
    klad=inv(c);
    chi_lr(i)=klad(1,2);
    m_ex=(-exp(-w)+exp(3*w))/(3*exp(-w)+exp(3*w));
    chi_ex(i)=m_ex-m_ex^2;
end;
subplot(2,2,1)
plot(w1,chi_lr,w1,chi_ex)
legend('\chi_{lr}','\chi_{ex}')
xlabel('w')
ylabel('\chi')
```

3. We use the recipe Eq. 45-48. From the data set we compute the means $m_1 = 1/3$ and $m_2 = -1/3$. The correlations in the data are given by $\langle s_1^2 \rangle_c = \langle s_2^2 \rangle_c = 1$ and $\langle s_1 s_2 \rangle_c = 1/3$. Thus the data covariance matrix is

$$C = \begin{pmatrix} 1 & 1/3 \\ 1/3 & 1 \end{pmatrix} - \begin{pmatrix} 1/9 & -1/9 \\ -1/9 & 1/9 \end{pmatrix} = \begin{pmatrix} 8/9 & 4/9 \\ 4/9 & 8/9 \end{pmatrix}$$

Its inverse is easily computed as

$$C^{-1} = \begin{pmatrix} 3/2 & -3/4 \\ -3/4 & 3/2 \end{pmatrix}$$

From Eq. 47 we obtain

$$w = \begin{pmatrix} 9/8 & 0 \\ 0 & 9/8 \end{pmatrix} - \begin{pmatrix} 3/2 & -3/4 \\ -3/4 & 3/2 \end{pmatrix} = \begin{pmatrix} -3/8 & 3/4 \\ 3/4 & -3/8 \end{pmatrix}$$

Finally,

$$\begin{aligned} \theta_1 &= \tanh^{-1}(m_1) + 3/8m_1 - 3/4m_2 = \tanh^{-1}(1/3) + 3/8 \\ \theta_1 &= \tanh^{-1}(m_2) - 3/4m_1 + 3/8m_2 = -\tanh^{-1}(1/3) - 3/8 \end{aligned}$$

Note, the appearance of the diagonal weights w_{ii}. In fact relation Eq. 36 ($\chi = C$) only holds off-diagonal. This relation is only equivalent to the relation $\chi^{-1} = C^{-1}$ if we also impose Eq. 36 for $i = j$. These are n additional relations and the solution can be obtained by introducing n additional parameters in the mean field description, which are the diagonal weights. This rather ad hoc argument can be made more formal using a higher order extension of the mean field theory using the Thouless Anderson Palmer (TAP) method. From this it is shown that the diagonal terms are always negative as we find in this example.