



Lecture 13 Quiz

[返回第 13 周课程](#)

7/7 得分 (100%)

测验通过！



1 / 1 分

1.

You may find this explanation of SBNs helpful.

This quiz is going to take you through the details of Sigmoid Belief Networks (SBNs). The most relevant videos are the second video ("Belief Nets", especially from 11:44) and third video ("Learning sigmoid belief nets") of lecture 13.

We'll be working with this network:

The network has no biases (or equivalently, the biases are always zero), so it has only two parameters: w_1 (the weight on the connection from h_1 to v) and w_2 (the weight on the connection from h_2 to v).

Remember, the units in an SBN are all binary, and the logistic function (also known as the sigmoid function) figures prominently in the definition of SBNs. These binary units, with their logistic/sigmoid probability function, are in a sense the stochastic equivalent of the deterministic logistic hidden units that we've seen often in earlier lectures.

Let's start with $w_1 = -6.90675478$ and $w_2 = 0.40546511$. These numbers were chosen to ensure that the answer to many questions is a very simple answer, which might make it easier to understand more of what's going on. Let's also pick a complete configuration to focus on: $h_1=0, h_2=1, v=1$ (we'll call that configuration C011).

Ready to Begin? (Please select a response. This question is reflective and selecting a certain answer will not affect your grade.)

☒ Yes

☐ No



1 / 1 分

2.

What is $P(v = 1 | h_1 = 0, h_2 = 1)$? Write your answer with four digits after the decimal point. Hint: the last three of

those four digits are zeros. (If you're lost on this question, then I strongly recommend that you do whatever you need to do to figure it out, before

proceeding with the rest of this quiz.)

0.6000

正确回答

Pretend that this is a feed-forward neural network with two hidden units and a logistic output neuron. You're now calculating the output of the network given that the hidden units have taken on the values $h_1 = 0$ and $h_2 = 1$.



1 / 1 分

3.

What is the probability of that full configuration, i.e.

$P(h_1 = 0, h_2 = 1, v = 1)$, which we called $P(C_{011})$? Write your answer with four digits

after the decimal point. Hint: it's less than a half, and the last two of those four digits are zeros.

0.1500

正确回答

We can use the rule of multiplication in order to obtain

$$P(h_1 = 0, h_2 = 1, v = 1) = P(v = 1 | h_1 = 0, h_2 = 1)P(h_1 = 0, h_2 = 1)$$

. Question 1 deals with finding $P(v_1 | h_1 = 0, h_2 = 1)$,

now you need to find $P(h_1 = 0, h_2 = 1)$. What does

the picture given in the preamble tell you about the

marginal independence of h_1 and h_2 (when we have not observed v)? Also, remember that h_1 and h_2 both have 0 total input, and that they are logistic neurons.



1 / 1 分

4.

What is $\frac{\partial \log P(C_{011})}{\partial w_1}$? Write your answer with at least three digits after the decimal point, and don't be too

surprised if it's a very simple answer.

0

正确答案



1 / 1 分

5.

What is $\frac{\partial \log P(C_{011})}{\partial w_2}$? Write your answer with at least three digits after the decimal point, and don't be too

surprised if it's a very simple answer.

0.4

正确答案



1 / 1 分

6. $w_1 = 10, w_2 = -4$

What is $P(h_2 = 1 | v = 1, h_1 = 0)$? Give your answer with at least four digits after the decimal point. Hint: it's a fairly small number (and not a

round number like for the earlier questions); try to

intuitively understand why it's small. Second hint: you might find Bayes' rule useful, but even with that rule, this still requires some thought.

0.0347

正确回答

We can use Baye's rule to determine:

$$P(h_2 = 1 | v = 1, h_1 = 0) = \frac{P(v=1|h_1=0, h_2=1)P(h_2=1)}{P(v=1|h_1=0, h_2=1)P(h_2=1) + P(v=1|h_1=0, h_2=0)P(h_2=0)}$$



1 / 1 分

7.

What is $P(h_2 = 1 | v = 1, h_1 = 1)$? Give your answer with at least four digits after the decimal point. Hint: it's quite different from the answer to

the previous question; try to understand why. The fact that those two are different shows that, conditional on the state of the visible units, the

hidden units have a strong effect on each other, i.e. they're not independent. That is what we call explaining away, and the earthquake vs. truck

network is another example of that.

0.4994

正确回答

