Activity: MLP Structure

ML for Health, Week 3

Part I.1

For each of the following, determine (a) how many logistic regression models and (b) how many parameters are contained in the model.

1. Logistic regression with 3 input features

    a) 1 logistic regression model
    b) <u>If we follow the convention of including 1 as an input feature</u>, then there are 3 parameters – one for each input feature
    c) <u>If we do not follow this convention</u>, then there are 4 parameters – one for each input feature, and one additional bias term

It may be helpful to draw or create graphs for these models. For this activity, bias/intercept parameters may be ignored.

Part I.2

For each of the following, determine (a) how many logistic regression models and (b) how many parameters are contained in the model.

2. An MLP with 3 input features and 1 hidden layer with 6 hidden units

a) There are 7 logistic regression models: one for each hidden unit, and 1 for the outcome
b) There are:
   i. 3 x 6 = 18 connections between the input features and the hidden units
   ii. 6 possible additional bias terms (one per hidden unit) depending on our convention (see Part I.1)
   iii. 6 connections between units in the hidden layer and the outcome
   iv. 1 bias term for the outcome
c) This gives us up to 31 parameters in total depending on handling of bias terms.

It may be helpful to draw or create graphs for these models. For this activity, bias/intercept parameters may be ignored.

Part I.3

For each of the following, determine (a) how many logistic regression models and (b) how many parameters are contained in the model.

3. An MLP with 3 input features and 2 hidden layers, each with 6 hidden units

a) There are 13 logistic regression models: 1 for each of the 6 * 2 = 12 hidden units, and 1 for the outcome

b) There are:

   i. 3 x 6 = 18 connections between the input features and the hidden units
   ii. 6 possible additional bias terms (one per hidden unit) depending on our convention (see Part I.1)
   iii. 6 x 6 = 36 connections between units in the first hidden layer and units in the second hidden layer
   iv. 6 bias terms, one per hidden unit in the second hidden layer
   v. 6 connections between units in the second hidden layer and the outcome
   vi. 1 bias term for the outcome

c) This gives us up to 73 parameters in total depending on handling of bias terms.

It may be helpful to draw or create graphs for these models. For this activity, bias/intercept parameters may be ignored.

Part I.4

For each of the following, determine (a) how many logistic regression models and (b) how many parameters are contained in the model.

4. (challenge) An MLP with 3 input features and 3 hidden layers with 6, 2, and 6 hidden units, respectively

a) There are 15 logistic regression models: 1 for each of the 6 + 2 + 6 = 14 hidden units, and 1 for the outcome

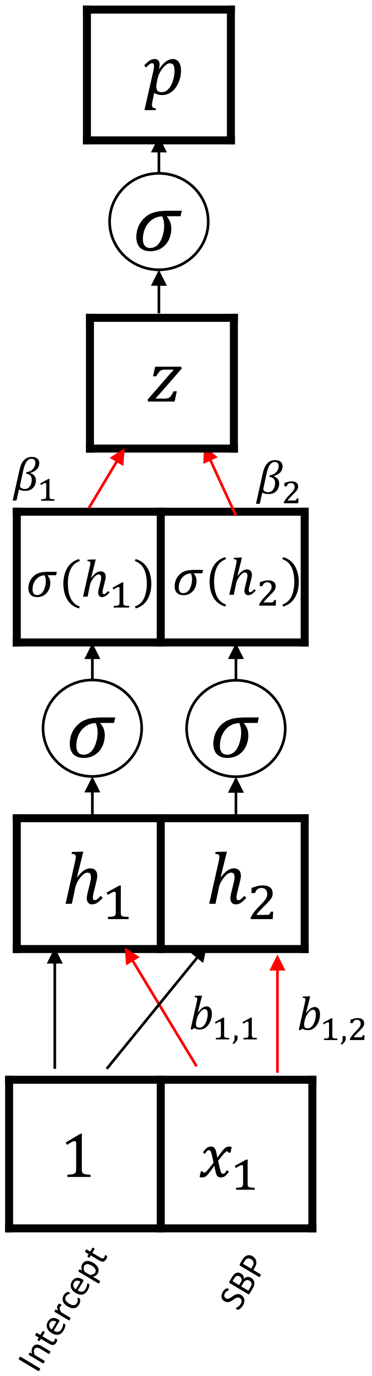
b) There are:
   i. 3 x 6 = 18 connections between the input features and the hidden units
   ii. 6 possible additional bias terms (one per hidden unit) depending on our convention (see Part I.1)
   iii. 6 x 2 = 12 connections between units in the first hidden layer and units in the second hidden layer
   iv. 2 bias terms, one per hidden unit in the second hidden layer
   v. 2 x 6 = 12 connections between units in the second hidden layer and units in the third hidden layer
   vi. 6 bias terms, one per hidden unit in the third hidden layer
   vii. 6 connections between units in the third hidden layer and the outcome
   viii. 1 bias term for the outcome

c) This gives us up to 63 parameters in total depending on handling of bias terms. This is fewer than the previous exercise even though this MLP has more hidden layers and more total hidden units!

It may be helpful to draw or create graphs for these models. For this activity, bias/intercept parameters may be ignored.

Part IIA

- The MLP at right is designed to predict ICU mortality from systolic blood pressure on admission.
- We would like to have a model that predicts high mortality risk associated with both *very high* AND *very low* systolic blood pressure.
- Let's suppose that h_1 detects *very high* blood pressure, h_2 detects *very low* blood pressure, and p is the model's final prediction about the probability of mortality.



Goal:

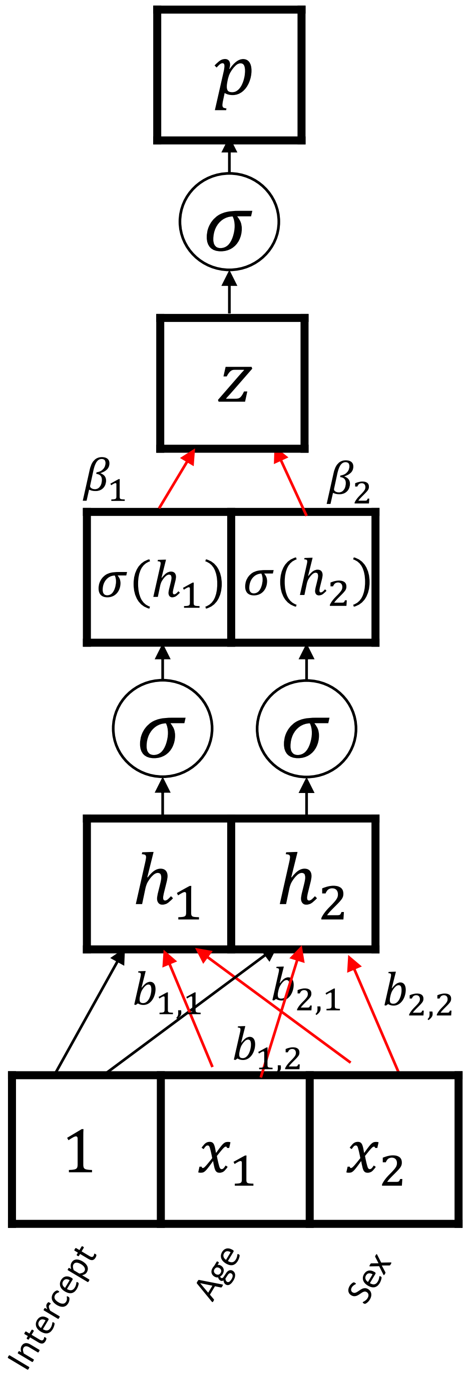
For each of the parameters highlighted in red, determine whether the value of that parameter should be (a) positive, or (b) negative.

Answers:

- $b_{1,1}$ must be positive: the log-odds of very high blood pressure increase as SBP increases
- $b_{1,2}$ must be negative: the log-odds of very low blood pressure decrease as SBP increases
- Both β_1 and β_2 must be positive, since the log-odds of mortality increases as (a) the probability of very high blood pressure increases, and (b) the probability of very low blood pressure increases
- Note that it is easier to understand relationships between predictors and coefficients if we consider *normalized* SBP values. After normalizing, we would have large negative SBP values for individuals with very low blood pressure and large positive SBP values for individuals with very high blood pressure

Part IIB

- The MLP at right is designed to predict disease mortality from age and sex
- We would like to have a model that predicts high mortality risk only for *males over 60* AND *females under 60*
- Let's suppose that h_1 detects *males over 60*, h_2 detects *females under 60*, and p is the model's final prediction about the probability of mortality.



Goal:

For each of the parameters highlighted in red, determine whether the value of that parameter should be (a) positive, or (b) negative.

Answers:

- $b_{1,1}$ must be positive: the log-odds of being a male over 60 increases with age
- $b_{1,2}$ must be negative: the log-odds of being a female under 60 decreases with age
- $b_{2,1}$ must be negative: the log-odds of being a male over 60 decreases with female sex ($x_2 = 1$) compared to male sex ($x_2 = 0$)
- $b_{2,2}$ must be positive: the log-odds of being a female under 60 increases with female sex ($x_2 = 1$) compared to male sex ($x_2 = 0$)
- Both β_1 and β_2 must be positive, since the log-odds of mortality increases as (a) the probability of being a male over 60 increases, and (b) the probability of being a female under 60 increases