LEARING FROM DATA

XIAOMO LIU

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

* *Department of Biology, University of Examples, London, United Kingdom*
[1] *Department of Chemistry, University of Examples, London, United Kingdom*

1 PERCEPTRON ALGORITHM

The perceptron algorithm is the most simple learning algorithm for classification problem. It works as follows

$$h(x) = \text{sign}\left(\left(\sum_{i=1}^{d} w_i x_i\right) + b\right) \tag{1}$$

where w_i is the weight for feature x_i and b is a bias term. We can easily rewrite it in terms of linear algebra. We will treat bias b as a constant weight $w_0 = b$. With this convention $\mathbf{w}^T\mathbf{x} = \sum_{i=0}^{d} w_i x_i$. From this notation, it looks like there is no difference between weights and bias in perceptron model. A natural question is that if weight and bias are similar, why we even borther to introduce both weight and bias?

2 LOGISTIC REGRESSION

The logistic regression utitilizes logistic function.

$$f(x) = \frac{L}{1 + e^{-k(x-x_0)}} \tag{2}$$

where L defines the curve's maximum value and k specifies steepness of the curve and x_0 is the centre of the curve. Sigmoid function is a special case of logistic function. The logistic regression is also a classification algorithm.

To answer this question, let take a look at the property of sigmoid function in Figure 1 and 2. In a nutshell, the weight w controls its steepness of sigmoid functions and bias b enables to shift the entire curve to left and right. Let take back and take glimpse at its propertity.

- Sigmoid has a nearly linear range around its centre 0 and becomes non-linear at two ends.
- Sigmoid has finite limits at negative infinity and infinity, most often going either from 0 to 1 or from -1 to 1 depending on convention.
- Sigmoid can model many natural processes, such as those of complex system learning curves, exhibit a progression from small beginings that accelerates and approaches a climax over time.
- Sigmoid is differeniable. Its derivative is $\frac{1}{dx}y = y(1-y)$.

The key difference between

3 METHODS

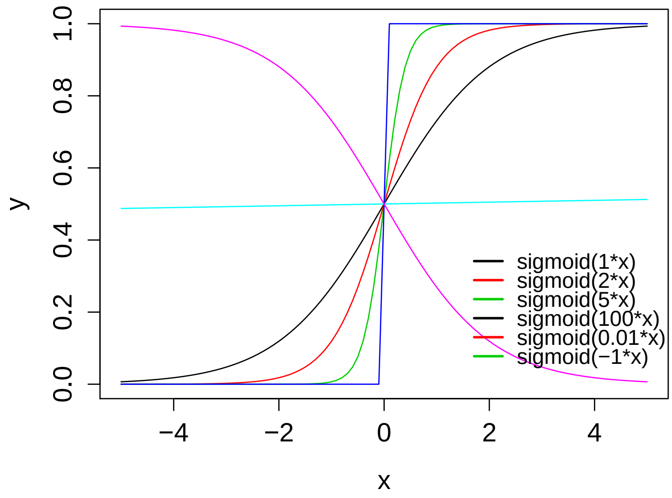

Figure 1: Sigmoid function with a single variant x, various choice of parameter w.

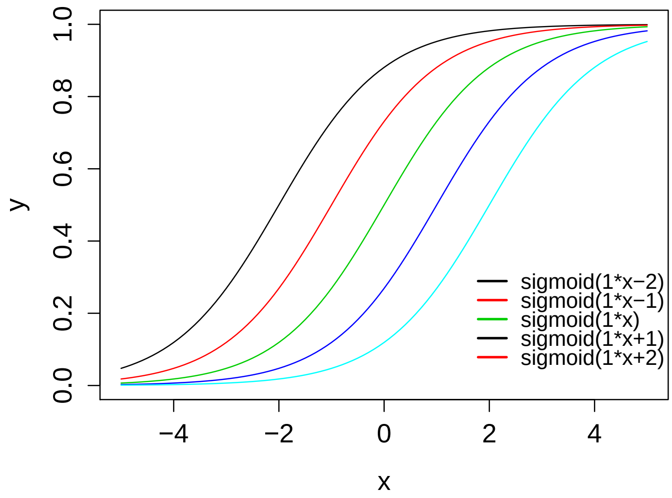

Figure 2: Sigmoid function with a single variant x, various choice of bias b.