Linear Data Analysis
Odds of Occurrence and Probability

Cain Susko

Queen's University
School of Computing

March 28, 2022

a Odds of Hyperplane Classification

Previously, We dealt with a hyperplane $\mathbb{H}$ and a vector $\vec{u}$ such that $\vec{u}$ is on one of 2 sides of the $\mathbb{H}$. This is useful but there is more information we can use to determine the location of $\vec{u}$ relative to $\mathbb{H}$. For example, the measurements of $\vec{u}$ are taken in a coordinate frame. Hyperplane $\mathbb{H}$ will have values in the same coordinate frame. This allows us to create a reference point on $\mathbb{H}$ called $\vec{r}$ and a corresponding unit vector $\vec{n}$; which allows us to determine the probability that $\vec{u}$ is on one side of $\mathbb{H}$ or the. This is to account for when points are very close to the hyperplane as the labeling is not as certain as the binary classification makes it seem.

Terms The definition of terms for probability in this course are:

Odds	ratio of likelihood of an event
Probability	inverse function of odds
Logistic Function	maps a score to a probability

We shall be dealing with the *Odds* that a data point will be in a given set.

b Odds and Probability

we shall first show by example what we mean by Odds. Given Odds (written as S), we say that S is the evaluation of the following ratios:

$$50:50 \qquad S=1$$

$$9:1 \qquad S=9$$

$$1:4 \qquad S=\frac{1}{4}$$

From these examples we can then infer that the Odds from a given ratio are:

$$S = \frac{p}{1-p}$$

where we are given $a : b$ such that $p = \frac{a}{a+b}$ which is the probability of the action occurring. This implies that S is greater than 0 with no upper bound.

Finding Probability Given the Odds S, can we find the probability p?

$$\begin{aligned} S &= \frac{p}{1-p} \\ &\equiv S(1-p) = p \\ &\equiv S - Sp = p \\ &\equiv S = p + Sp \\ &\equiv S = p(1+S) \\ &\equiv p = \frac{S}{S+1} \end{aligned}$$

Thus, the formula for finding the probability of an outcome given it's odds is:

$$p = \frac{S}{1+S}$$

c Logistic Function for Odds

this section relates the score of a data point to it's probability using the Logistic Function. The bounds of an Odd and a Score are:

Odd	$S \in (0, +\infty)$
Score	$z \in (-\infty, +\infty)$

The Logistic function should map an Odd to an Score such that:

$$f : S \mapsto z$$

where the resulting graph is invertible, smooth, 1:1, and monotone. These constrains allow for the score to equal:

$$z = ln(S)$$

such that:

$$S = e^z$$

Now, given a hyperplane $\mathbb{H}$, a unit normal vector $\vec{n}$ and reference point a (same as weights vector and bias scalar for $\mathbb{H}$) . The score for a given value $\vec{u}$ is equal to:

$$z(\vec{u}) = \vec{n}^\top \vec{u} + a$$

Now, to relate all these inferences, the way we relate a probability to a score is:

$$p(z) = \frac{e^z}{1 + e^z}$$

$$= \frac{1}{1 + e^{-z}}$$

A consequence of this equation is that if z is negative, then:

$$p(-z) = 1 - p(z)$$

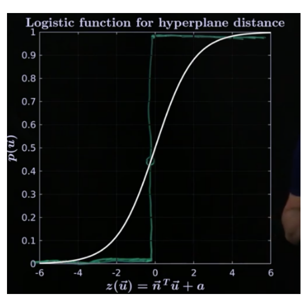
Thus the Odds of $\vec{u}$ of being in the **positive half-space** of $\mathbb{H}$ is:

$$z(\vec{u}) = \vec{n}^\top \vec{u} + a$$

$$p(\vec{u}; \vec{n}, a) = \frac{1}{1 + e^{-(\vec{n}^\top \vec{u} + a)}}$$

d Properties of The Logistic Function

given the scores and probability of data-vectors from a logistic function we can plot the following graph:



Where the white line is the Logistic function and the green line is the binary classification function. Additionally the x axis is the score and the y axis is the probability for $\vec{u}$.

Derivative of p(z)

The equation for relating score to a probability is:

$$p(z) = \frac{e^z}{1 + e^z}$$

observe:

$$\begin{aligned} 1 - p(z) &= \frac{1 + e^z}{1 + e^z} - \frac{e^z}{1 + e^z} \\ &= \frac{1}{1 + e^z} \\ \\ &= p(-z) \end{aligned}$$

Recall The derivative of a fraction is:

$$\frac{d}{dz}\left(\frac{f(z)}{g(z)}\right) = \frac{f'g - fg'}{g^2}$$

Additionally, the derivative of e^x is:

$$\frac{d}{dz}e^z = e^z$$

Solution The formula for the derivative of $p(z)$ is:

$$\begin{aligned} \frac{d}{dz}p(z) &= \frac{e^z(1 + e^z) - e^z e^z}{(1 + e^z)^2} \\ \\ &= p(z)(1 - p(z)) \end{aligned}$$

Summary

Probability

- Odds: ratio involving probability
- Probability: function of distance from hyperplane

Logistic Function

- continuous, differentiable, invertible

Learning Outcomes

Students should now be able to:

- Transform between Odds and Probability
- Find unit form of hyperplane $\mathbb{H}$
- find probability that $\vec{u}$ is on positive side of $\mathbb{H}$