Logistic Regression

Binary Classification

- X is one or more attributes
- y is in {0,1}
- Example
  - x - size of tumor
  - y - malignant? (yes or no)
  - $\hat{y}$ - predicted value of y (sometimes y')

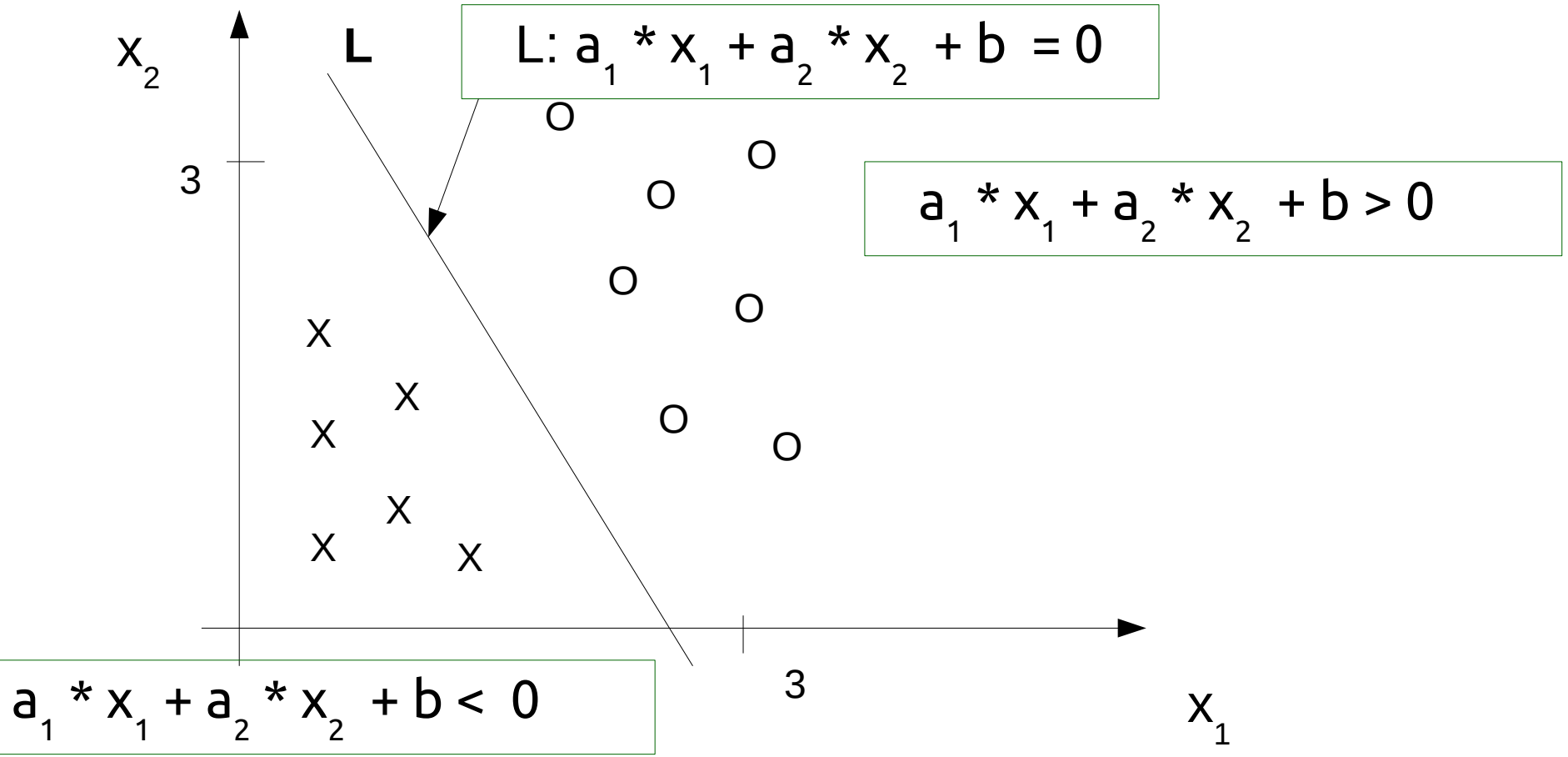
Classifiers

- Probabilistic Classifier – outputs probabilities of an instance belonging to a particular class.
- For binary classification ouput the probability p of a positive.
- Discrete classifier – outputs a predicted class rather than a probability.

Probabilistic Classifier

- Logistic Regression builds a Probalistic Classifier.
- So $\hat{y}$ (or p) is between 0 and 1.
- $\hat{y}$ is the probability of a positive.
- (A positive is an event of interest, for example having a medical condition.)

Example - Two Features



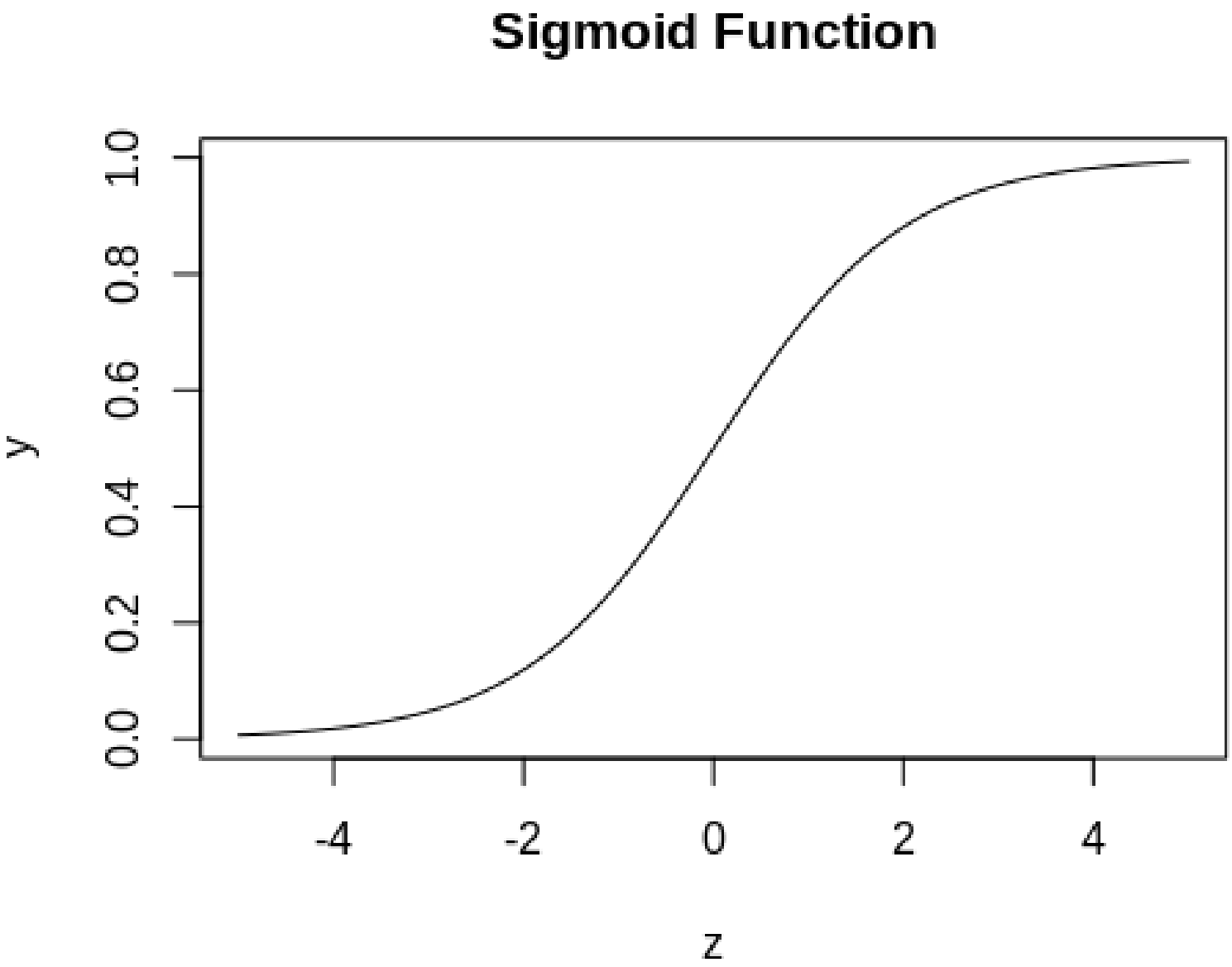
Separation Line

- Suppose for a moment we have the value of the parameters a, b_1 and b_2.
- A value of $a_1 * x_1 + a_2 * x_2 + b < 0$ indicates an x.
- A value of $a_1 * x_1 + a_2 * x_2 + b > 0$ indicates an o.
- However the output now is between $-\infty$ and ∞.
- What we need is a value between 0 and 1.

Sigmoid Function

$$y = \frac{1}{\left(1 + e^{-z}\right)}$$

- Also called the logistic function
- Maps all z values to values between 0 and 1
- Its a way to maps values between -∞ and ∞ to values between 0 and 1.


Sigmoid Function
y
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
z

Logistic Regression

- Want $0 <= \hat{y} <= 1$
- $z = a_1 * x_1 + .. a_n * x_n + b$

$$z = \sum_k a_k x_k + b$$

- z can take values between $-\infty$ and ∞
- Apply a function g(z) to z to give values between 0 and 1.

$$\hat{y} = g(z) = \frac{1}{(1 + e^{-z})}$$

Interpretation

- Can now interpret $\hat{y}$ as a probability of a positive value (1).
- For example, $\hat{y} = 0.7$
  - 70% chance of tumor being malignant.
  - 30% change of it not being malignant.

Prediction/ Threshold

- In order to make a prediction
  - choose threshold (say 0.5)
  - if $\hat{y}$ > threshold, predict y = 1.
- [Notice that $\hat{y} = g(z) > 0.5$ when $z > 0$]
- Thresholds will be examined in detail in the next section.
- A threshold converts a probabilistic classifier to discrete classifier.