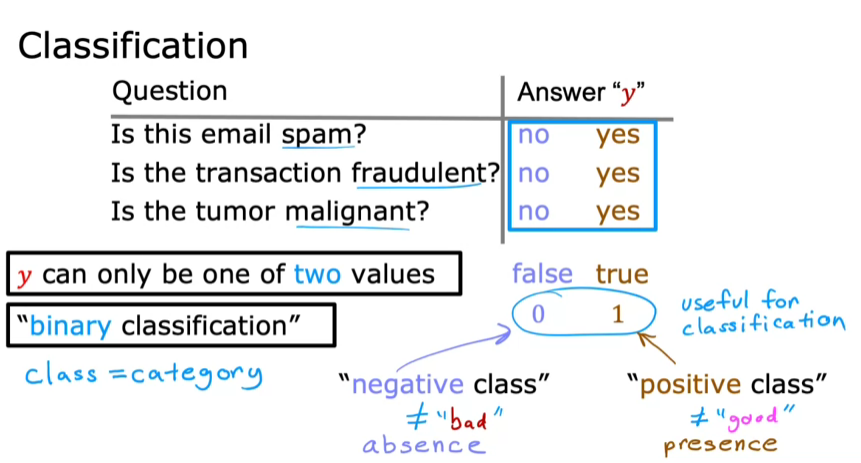
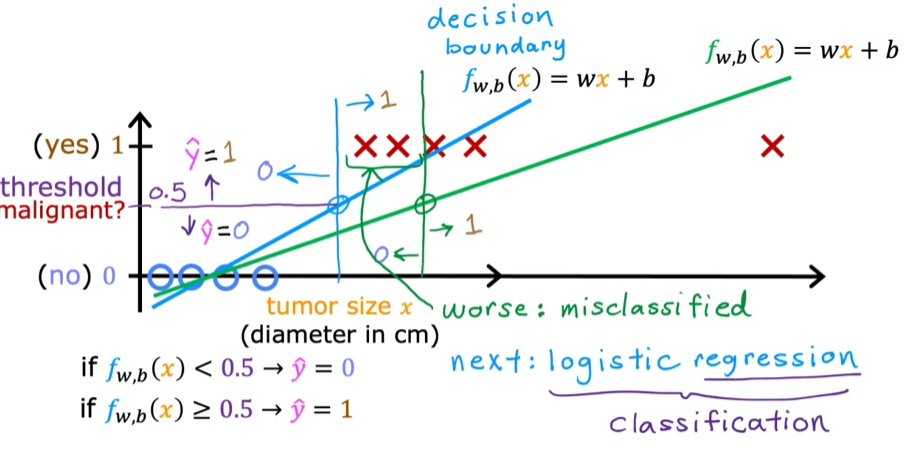
**CLASSIFICATION**

**MOTIVATION**

* **Classification predicts a limited set of outcomes, such as determining if an email is spam (yes or no) or if a financial transaction is fraudulent (true or false).**
* **Binary classification involves two possible outputs, often represented as 0 (negative class) and 1 (positive class).**

****

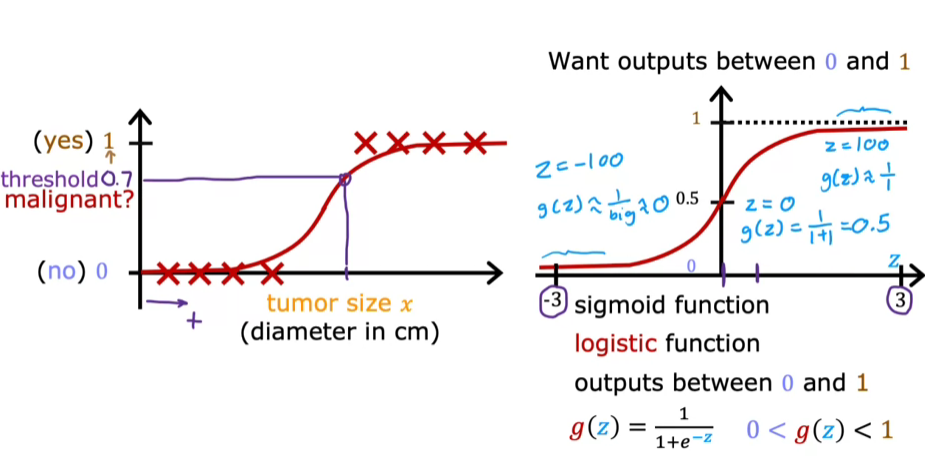
* **Linear regression predicts a continuous range of values, which is not ideal for classification tasks where outputs should be discrete categories.**
* **Adding new data points can shift the decision boundary inappropriately, leading to incorrect classifications.**

****

* **Logistic regression is designed for binary classification, providing outputs that are constrained between 0 and 1, which helps avoid the issues seen with linear regression.**
* **Despite its name, logistic regression is primarily used for classification tasks, not regression, which can be confusing but is important to understand.**

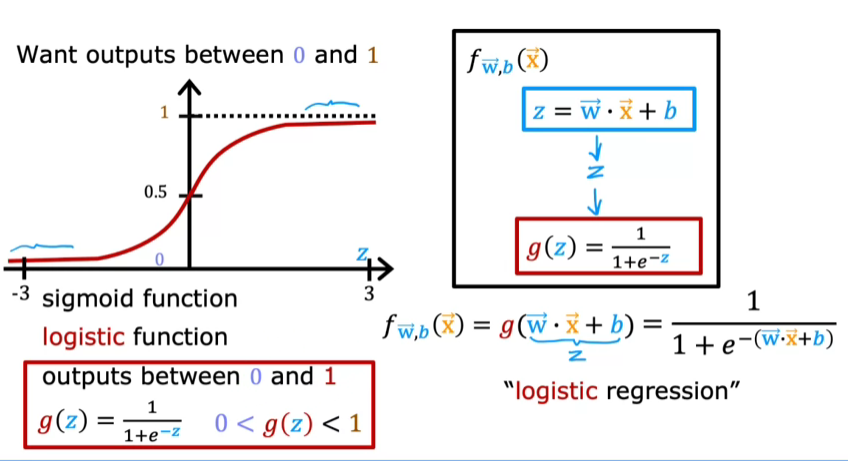
**LOGISTIC REGRESSION**

* **Logistic regression uses labels 1 (malignant) and 0 (benign) to classify tumors based on their size, represented in a graph where the output is either 0 or 1.**
* **Unlike linear regression, logistic regression fits an S-shaped curve to the data, allowing for probabilities between 0 and 1.**

****

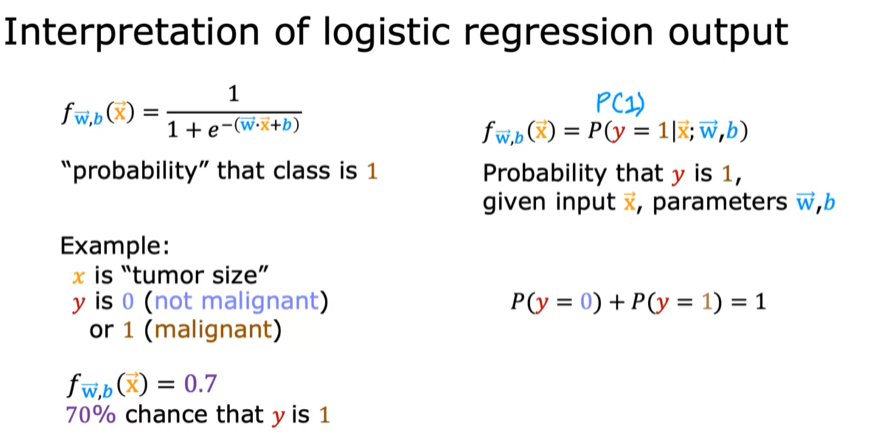
**The Sigmoid function**

* **The Sigmoid function, or logistic function, is defined mathematically as g(z) = 1 / (1 + e^(-z)), where e is approximately 2.7.**
* **This function outputs values between 0 and 1, with specific behavior at extreme values of z, making it suitable for classification tasks.**

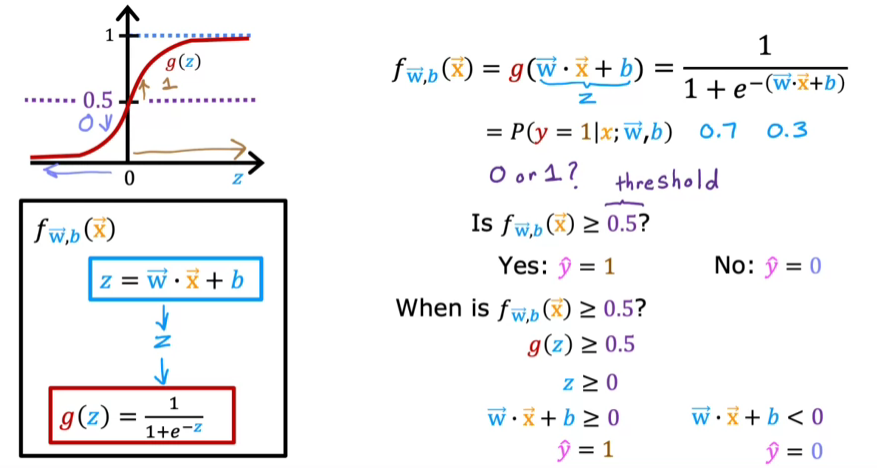
****

**Interpreting logistic regression output**

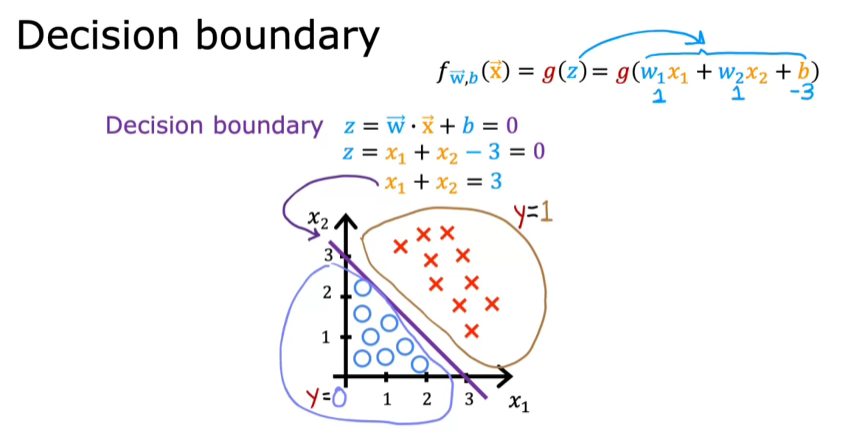
* **The output of logistic regression represents the probability that a given input belongs to the positive class (y = 1). For example, a probability of 0.7 indicates a 70% chance of malignancy.**
* **The probabilities for y = 0 and y = 1 must sum to 1, so if y = 1 has a 70% chance, then y = 0 has a 30% chance.**

****

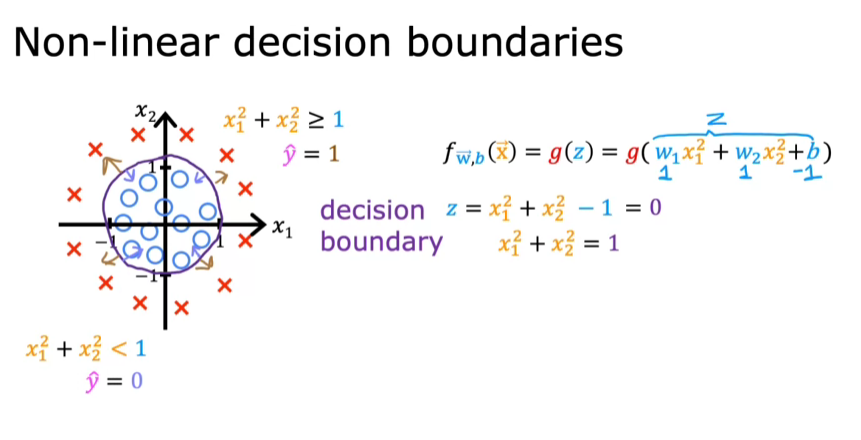
**DECISION BOUNDARY**

****

* **Predictions are made in two steps: first, compute (z = w. x + b), then apply the Sigmoid function (g(z)) to obtain the probability (f(x)).**
* **A common threshold of 0.5 is used to classify predictions: if (f(x) >= 0.5), predict (y = 1); if (f(x) < 0.5), predict (y = 0).**

****

* **The decision boundary is defined by the line where (z = 0) (i.e., (w. x + b = 0)), separating the predictions of (y = 1) and (y = 0).**
* **For example, with parameters (w1 = 1, w2 = 1, b = -3), the decision boundary is the line (x1 + x2 = 3).**

****

* **By incorporating polynomial features, logistic regression can model more complex decision boundaries, such as circles or ellipses.**
* **Higher-order polynomial terms allow for even more intricate decision boundaries, enabling the model to fit complex data patterns.**