Linear Classifiers (Part 2)

CS114B Lab 3

Kenneth Lai

February 10, 2022

Linear Classifiers

- What is g?

Linear Classifiers

- What is g?
  - Function with many names
    - In the context of generalized linear models, this is called a link function
    - In the context of neural networks, this is called an activation function

Linear Classifiers

- What is g?
  - Function with many names
    - In the context of generalized linear models, this is called a link function
    - In the context of neural networks, this is called an activation function
- A dot product of two vectors produces a scalar, but in general, we don't just want an arbitrary real number

Linear Classifiers

- What is g?
  - Function with many names
    - In the context of generalized linear models, this is called a link function
    - In the context of neural networks, this is called an activation function
- A dot product of two vectors produces a scalar, but in general, we don't just want an arbitrary real number
  - Sometimes, we want a probability (logistic regression)
  - Sometimes, we just want the decision itself (perceptron)

Linear Classifiers

- Let $z = \theta \cdot \mathbf{x}$ (or $\mathbf{w} \cdot \mathbf{x} + b$)

Linear Classifiers

- Let $z = \theta \cdot \mathbf{x}$ (or $\mathbf{w} \cdot \mathbf{x} + b$)
- Logistic regression: logistic (sigmoid) function
  - $\sigma(z) = \dfrac{1}{1 + e^{-z}}$

Linear Classifiers

- Let $z = \theta \cdot \mathbf{x}$ (or $\mathbf{w} \cdot \mathbf{x} + b$)
- Logistic regression: logistic (sigmoid) function
  - $\sigma(z) = \dfrac{1}{1 + e^{-z}}$
- Perceptron: (Heaviside) step function
  - $H(z) = \begin{cases} 1, & \text{if } z > 0 \\ 0, & \text{if } z < 0 \end{cases}$

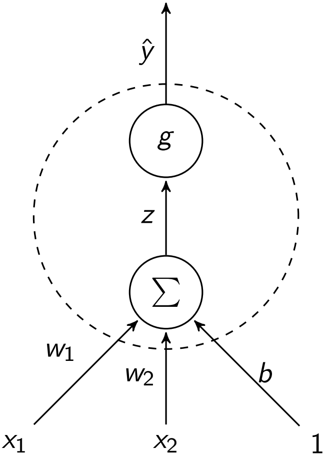
Linear Classifiers

- Let $z = \theta \cdot \mathbf{x}$ (or $\mathbf{w} \cdot \mathbf{x} + b$)
- Logistic regression: logistic (sigmoid) function
  - $\sigma(z) = \dfrac{1}{1 + e^{-z}}$
- Perceptron: (Heaviside) step function
  - $H(z) = \begin{cases} 1, & \text{if } z > 0 \\ 0, & \text{if } z < 0 \end{cases}$
  - What if $z = 0$?

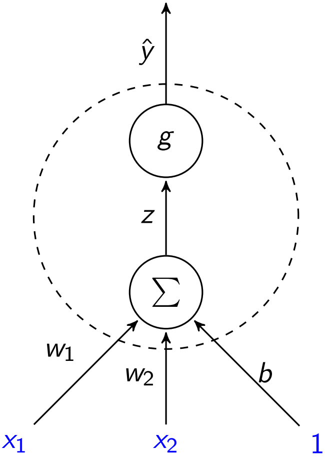
Linear Classifiers

- Let $z = \theta \cdot \mathbf{x}$ (or $\mathbf{w} \cdot \mathbf{x} + b$)
- Logistic regression: logistic (sigmoid) function
  - $\sigma(z) = \dfrac{1}{1 + e^{-z}}$
- Perceptron: (Heaviside) step function
  - $H(z) = \begin{cases} 1, & \text{if } z > 0 \\ 0, & \text{if } z < 0 \end{cases}$
  - What if $z = 0$?
    - Set by convention (1, 0, or 1/2)

Graphical Representation of a Linear Classifier (1)

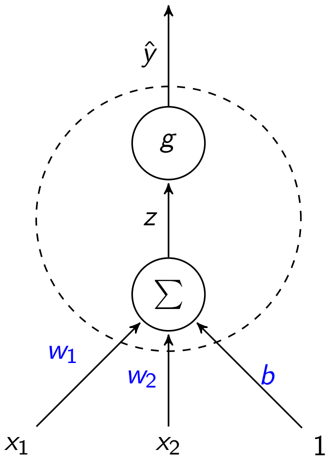


Graphical Representation of a Linear Classifier (1)



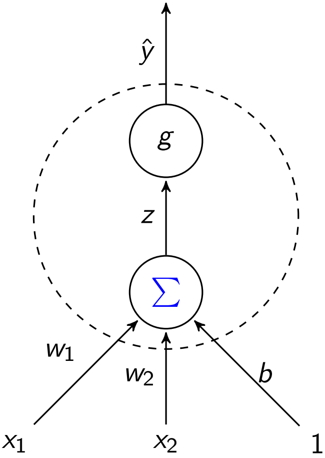
- Input (including dummy feature 1)

Graphical Representation of a Linear Classifier (1)



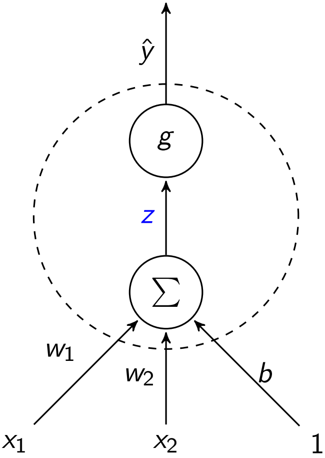
- Parameters (weights and bias term)

Graphical Representation of a Linear Classifier (1)



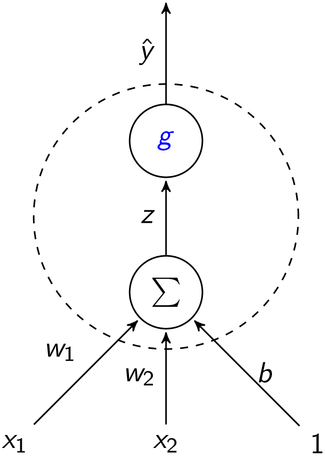
- Sum function $\sum$

Graphical Representation of a Linear Classifier (1)



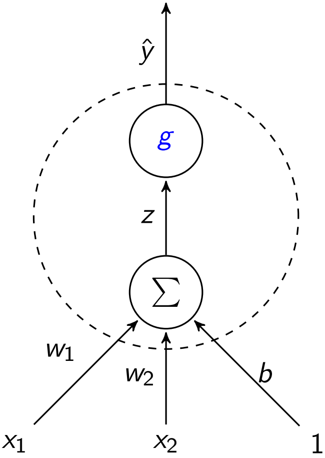
- "Score"

Graphical Representation of a Linear Classifier (1)



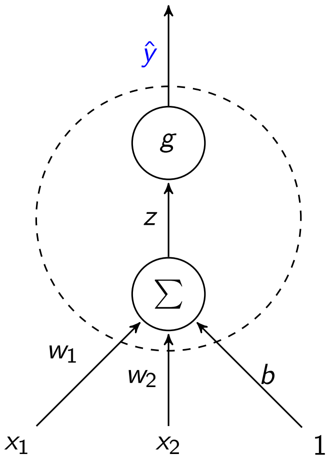
- Activation function g

Graphical Representation of a Linear Classifier (1)



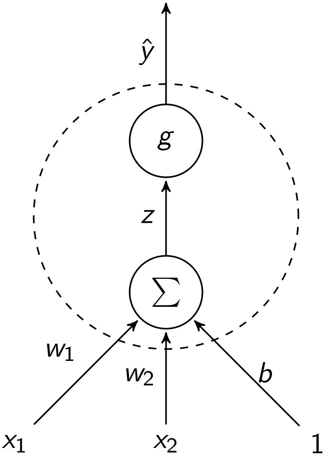
- Activation function g
  - Logistic, step, etc.

Graphical Representation of a Linear Classifier (1)

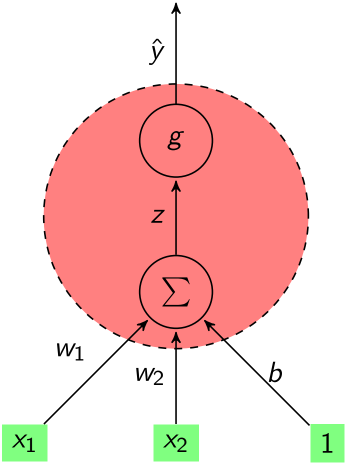


- Output

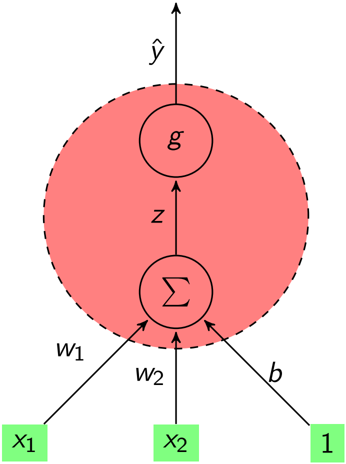
Graphical Representation of a Linear Classifier (1)



Graphical Representation of a Linear Classifier (1)

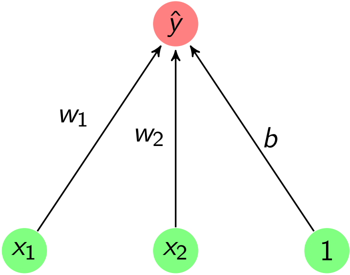


Graphical Representation of a Linear Classifier (1)



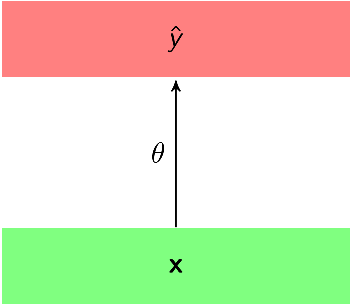
- Output
- Input

Graphical Representation of a Linear Classifier (1)



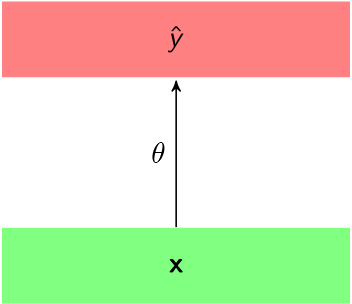
- Output
- Input

Graphical Representation of a Linear Classifier (2)



- Output
- Input

Graphical Representation of a Linear Classifier (2)



- Output
- Input
- $\hat{y} = g(\theta \cdot \mathbf{x})$
  - We will assume that the dummy feature 1 is part of $\mathbf{x}$