SGN-13006 Introduction to Pattern Recognition and Machine Learning (5 cr)

Linear regression and classification

Joni-Kristian Kämäräinen

August 2018

Laboratory of Signal Processing
Tampere University of Technology

Material

- Lecturer's slides and blackboard notes
- Very thorough treatment in C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006: e.g. Chapter 4

Linear models are at the core of many powerful methods in PR and ML and therefore can be found from all ML & PR books - they provide intuitive mathematical introduction to the topic

Linear regression

Linear regression using high school maths

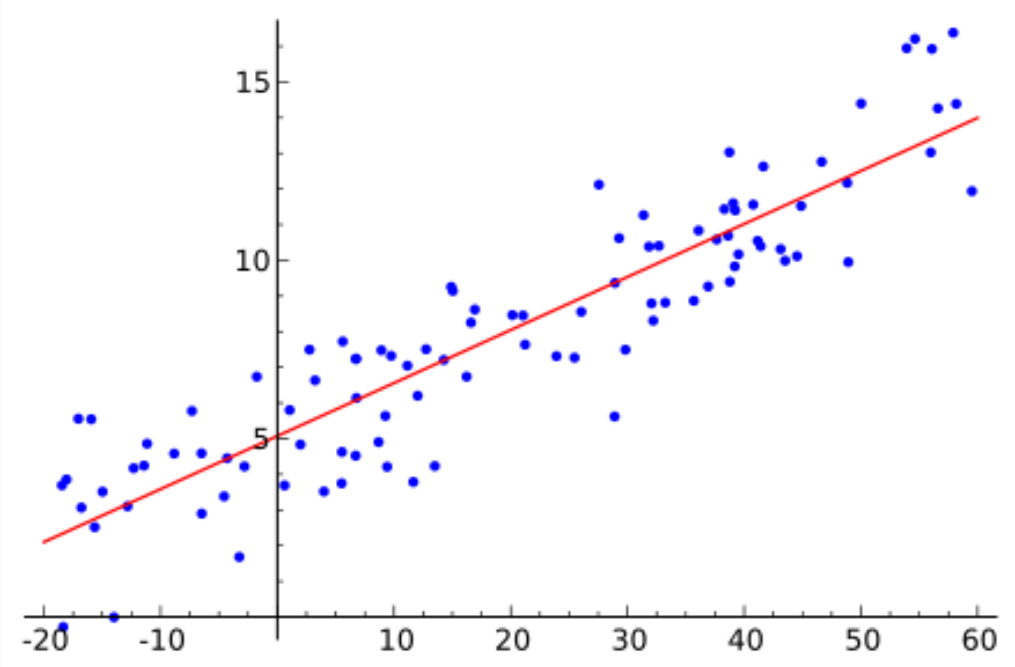

Figure 1: Linear regression (Wikipedia).

Linear classification

Linear classification

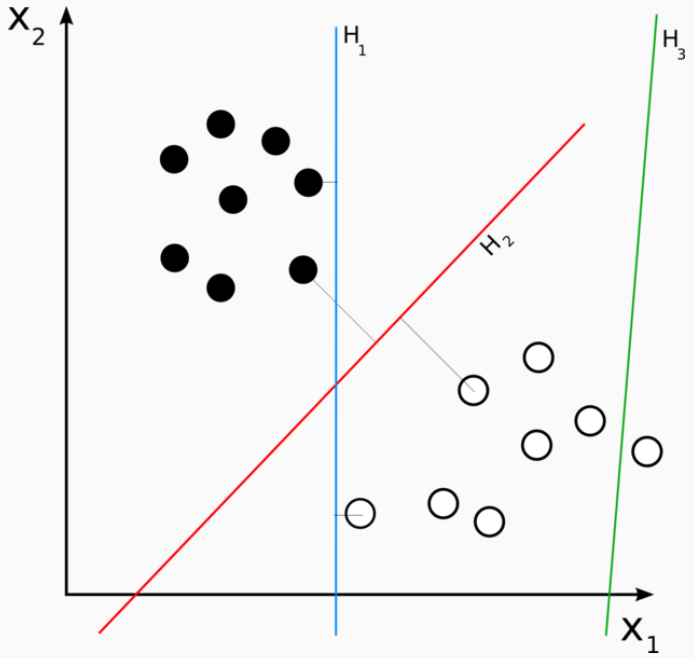

Figure 2: Linear classification (Wikipedia).

Linear regression and classification in PR and ML

Linear regression and classification in PR and ML

- The two essential problems in PR and ML are *regression* and *classification*.
- We derived solutions to these problems by using high school geometry (fitting lines/hyperplanes to a set of points) and simple algebra $\rightarrow$ This magic is not that difficult after all!
- PR and ML are about different models of learning (e.g. $y = ax + b$) and different methods of learning using the models (e.g. LSQFit of a, b).
- However, He-Who-Must-Not-Be-Named is in the details and true mastery is to avoid His traps!

Summary

Summary

Summary

1. Linear regression: Fitting a line (hyperplane) to points (1, 2, . . . , N points in 2D) of input x to predict output y

Summary

1. Linear regression: Fitting a line (hyperplane) to points (1, 2, . . . , N points in 2D) of input x to predict output y
2. Linear classification: Fitting a line (hyperplane) to points from two different classes (1D and 2D)

Summary

1. Linear regression: Fitting a line (hyperplane) to points (1, 2, . . . , N points in 2D) of input x to predict output y
2. Linear classification: Fitting a line (hyperplane) to points from two different classes (1D and 2D)
3. What can make line fitting to fail?
    - Too few data points
    - Outliers
    - Division-by-zero (singularities)
    - Computational complexity (time/space)
    - Multiple solutions (local vs global optima)
    - Model selection (low vs. high order polynomial)
    - Background (prior) knowledge