University of Zurich[UZH]

Institute of Computational Linguistics

Introduction to Machine Learning

Lesson 3

Mathias Müller, Phillip Ströbel

True or false?

A pairplot shows the relationship between all features and the target variable.

K Nearest Neighbour is fast to train, and slow during inference.

The k in K Nearest Neighbour is the number of classes in the training data.

University of Zurich[UZH]

Institute of Computational Linguistics

Linear Regression

Topic of this lesson

- Linear regression
- Evaluation metrics for linear regression
- (Polynomial Features for linear regression)

Regression Problems

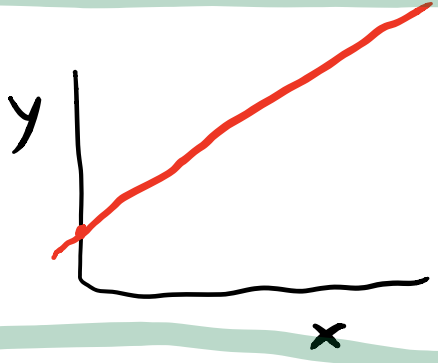
- Assumption: data-generating process is a function
- fitting a regression model: approximating this unknown function

- fitting a regression model: 1) decide on a class of functions, 2) set all parameters that fully describe the function

Classes of functions

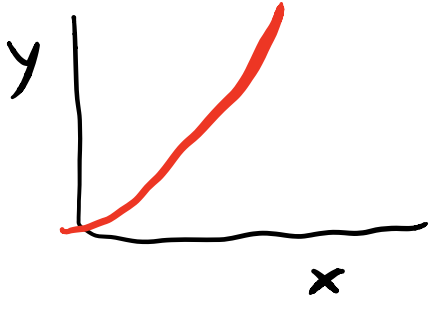
linear functions

$$y = 2x + 1$$



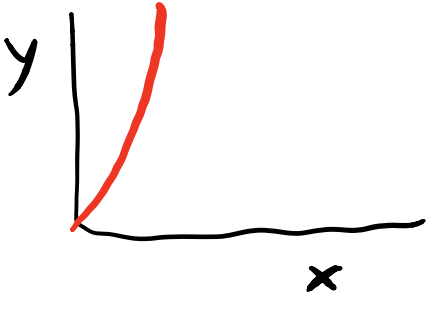
polynomial functions

$$y = 3x^2 + 2x$$



exponential functions

$$y = e^x$$



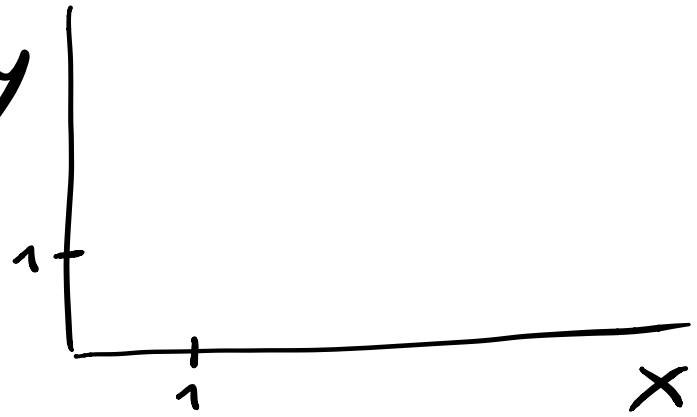
Parameters that describe functions

$+32$

y intercept

$y = -2x + 3$

Coefficient



Linear Regression

- function class: **linear**
- linear functions describe **lines or hyperplanes**
- parameters to be learned: 1 weight for each feature in X, optionally 1 **intercept**

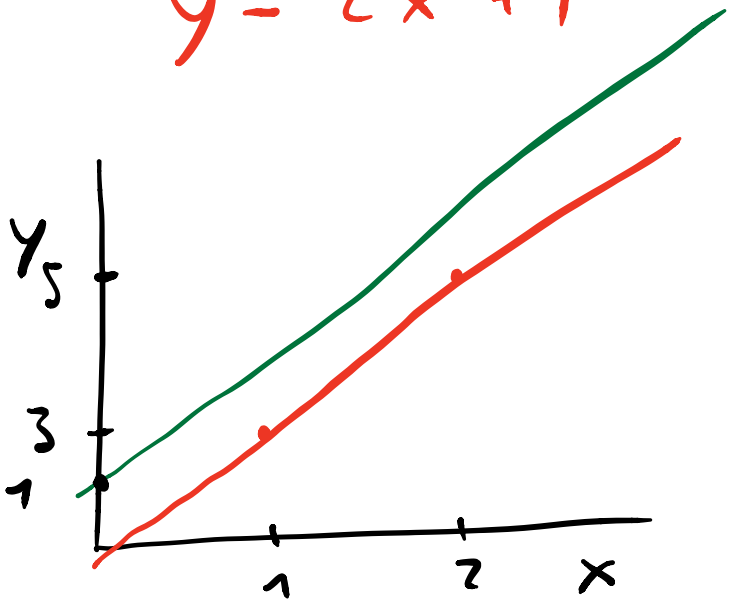
$$y = 3x_1 + 7x_2 + 1$$

Line or Hyperplane?

$x = 1 \quad y = 3$

$x = 2 \quad y = 5$

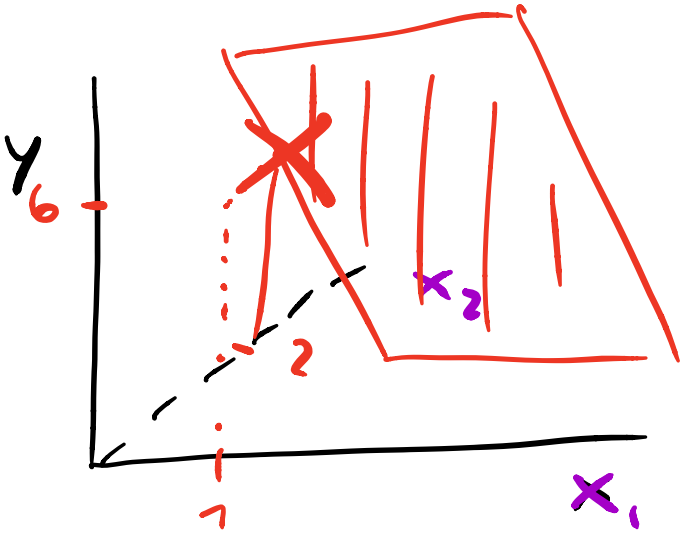
$y = 2x + 1$



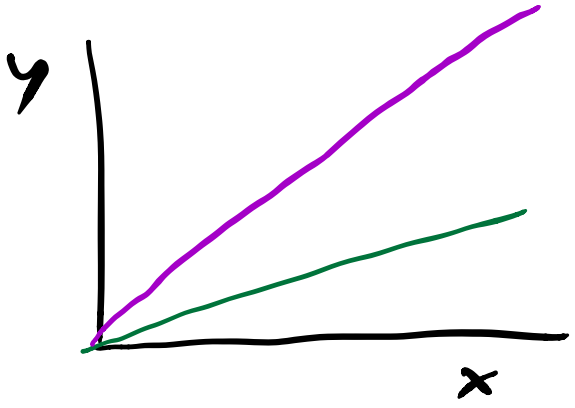
$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}^{x_1}_{x_2} \quad y = 6$

$x = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad y = 8$

$y =$



Equation of a line



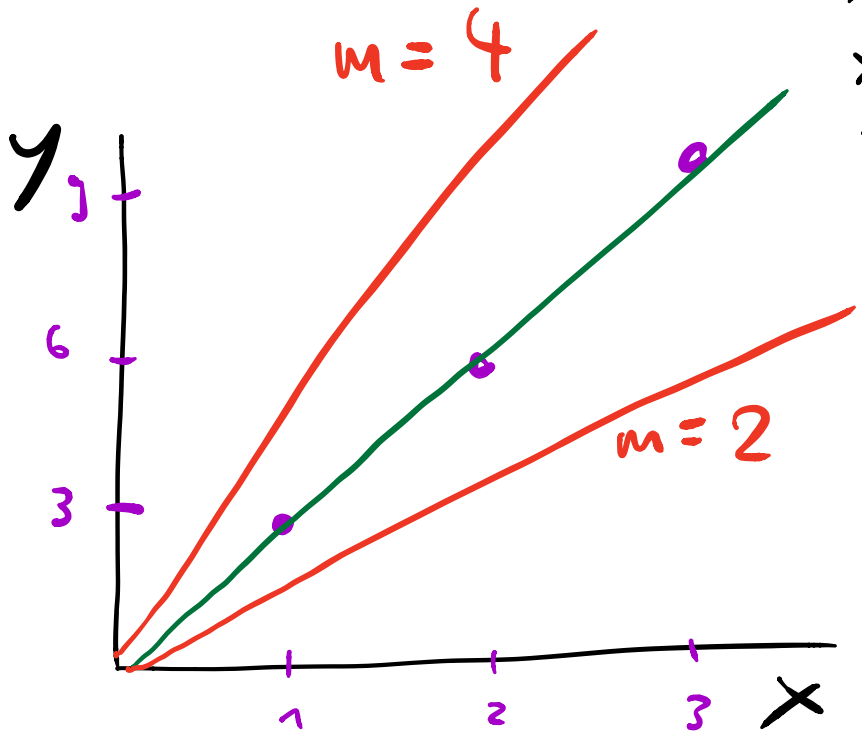
$y = mx + b$

Steigung

$y = 2x$

$y = 13x$

Simple linear regression problem: one feature, one target variable, no intercept



$x = 1$ $y = 3$
$x = 2$ $y = 6$
$x = 3$ $y = 9$

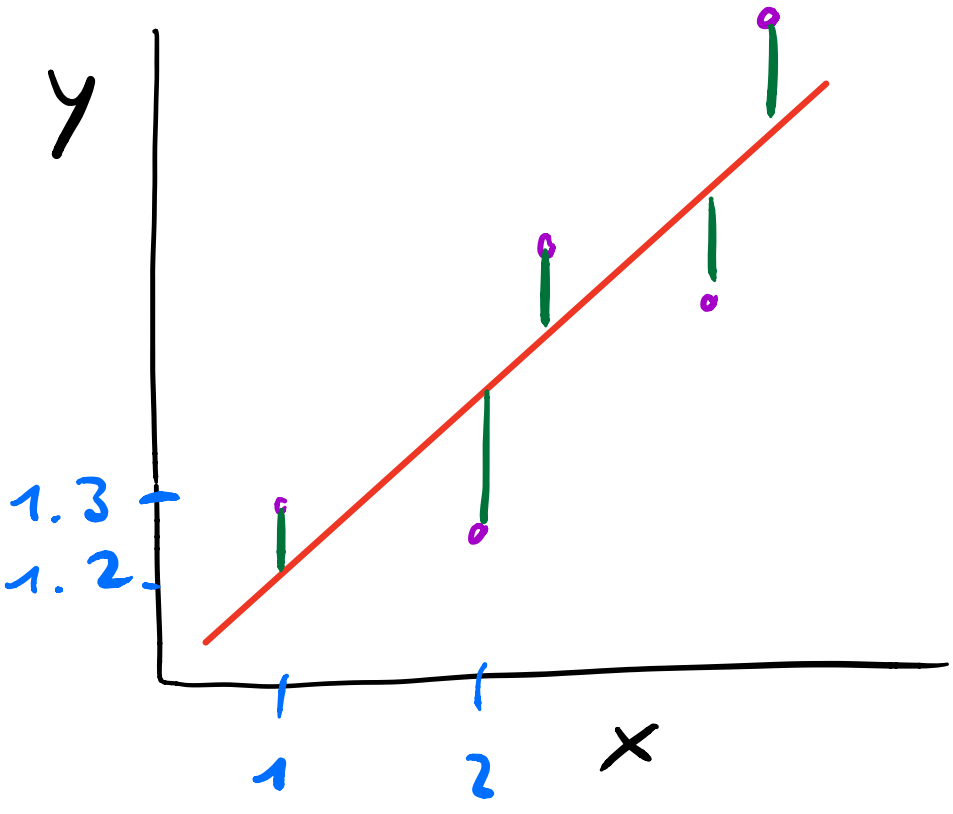
training data

$y = cx$

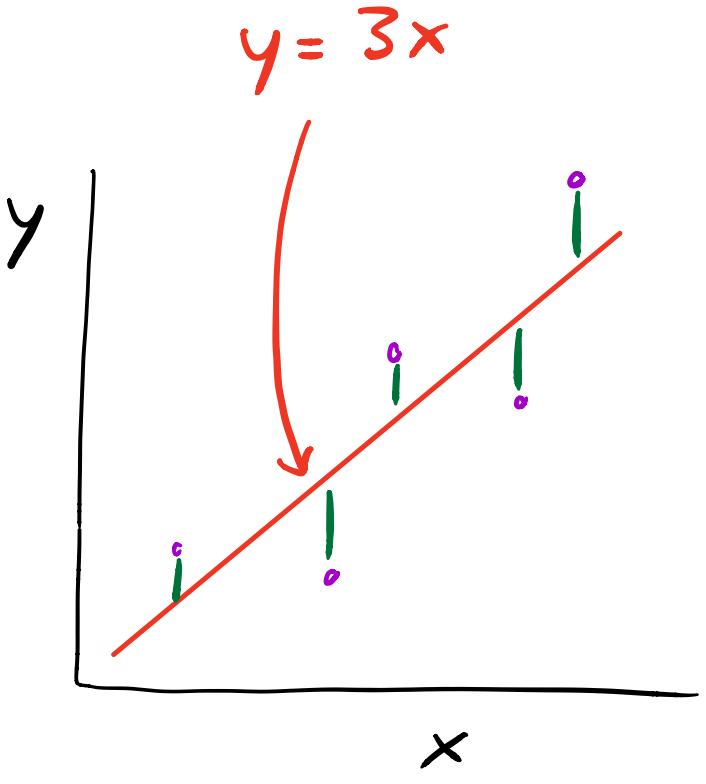
$c = 3$

Residuals

o = training sample

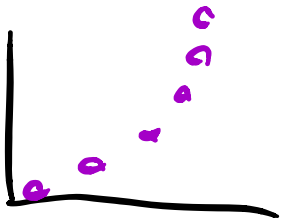


Goodness of fit: sum of squared residuals

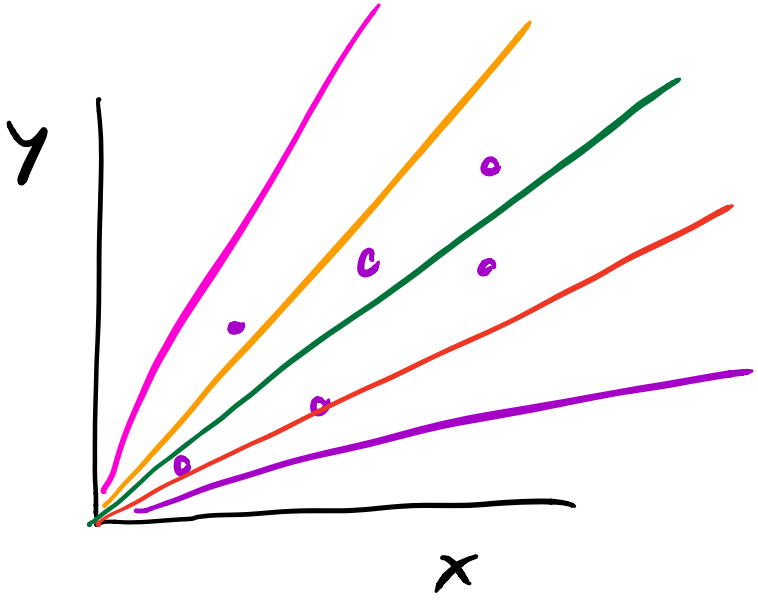


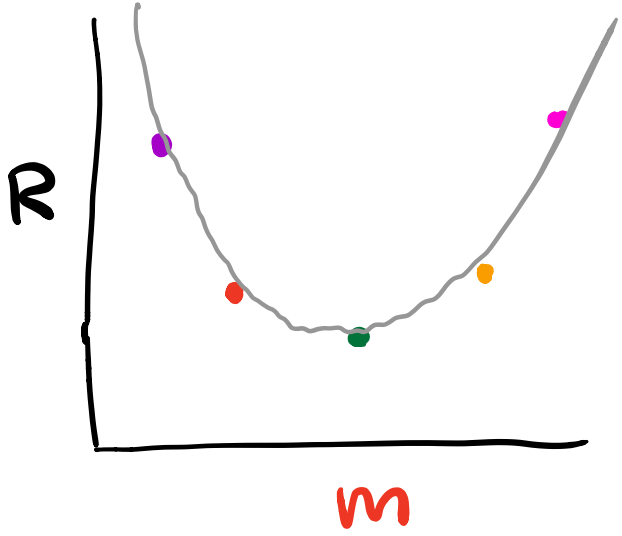
$$R = \sum_{x,y} (y - f(x))^2$$

How to find best line? Let's analyze sum of squared residuals



$y = mx$





ordinary least squares

Ordinary least squares (OLS)

- closed form, analytical solution for linear functions

Summary

- Regression approximates functions that generated the data
- functions are defined by their parameters

- linear regression approximates linear functions
- linear functions are lines or hyperplanes

- model fitting means finding parameters that minimize sum of squared residuals, with OLS
- Polynomial features to fit non-linear data
- metrics for regression problems: