AGH University of Science and Technology

Seminar in *Artificial Intelligence*

Marcin Zajac, Dominik Koza, Lukasz Gorczyca

Department of Telecommunications

01.04.2019

Agenda

- 1. Intro, presentation plan.
- 2. What is regression?
- 3. What regression is used for?
- 4. Types of regression.
- 5. Linear regression.
- 6. Logistic regressiion.
- 7. Polynominal regression.
- 8. QA
- 9. Quiz.

2. What is regression?

3. What regression is used for?

4. Types of regression

- linear regression
- logistic regression
- polynominal regression
- stepwise regression
- ridge regression
- lasso regression
- elasticNet regression

5.Linear regression

- First known research in this area - method of least squares published by Legendre in 1805 and by Gauss in 1809
- The representation is a linear equation that combines a specific set of input values x the solution to which is the predicted output for that set of input values y. As such, both the input values x and the output value are numeric.

5.Simple Linear Regression

- Simple linear regression is a linear regression model with a single independent variable
- Model for single dimension

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{1}$$

- Naming:
  - The unknown parameters - β
  - The independent variables - X or x
  - The dependent variable - Y or y
  - Introduced error - ϵ

5.Multiple dimension extension.

- When there is a single input variable x, the method is referred to as simple linear regression. When there are multiple input variables, literature from statistics often refers to the method as multiple linear regression.
- Model for n dimension

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + ... + \beta_n x_{i,n} + \epsilon_i \tag{2}$$

- To matrix representation

$$y_i = x_i^T + \epsilon_i \tag{3}$$

$$Y = X\beta + \epsilon \tag{4}$$

5.Ordinary least squares

$$\hat{\beta} = argmin_{\beta} S(\beta) \tag{5}$$

$$S(\beta) = \sum_{i=1}^{n} \mid y_i - \sum_{j=1}^{p} X_{ij}\beta_j \mid^2 \tag{6}$$

Gradient descent

Regularization

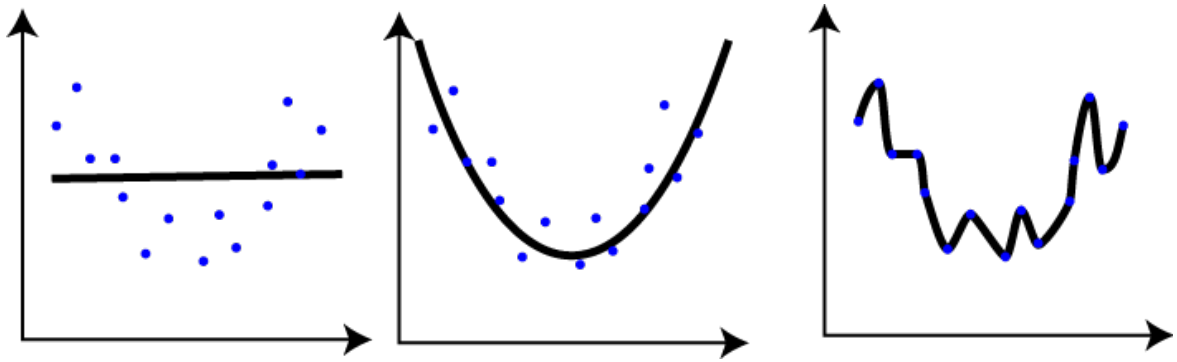
6. Logistic regression

7.Polynominal Regression

- Occurs when regression equation has independent variable in power higher than 1.
- Example (variable in 2 power)

$$y = a + b * x^2 \tag{7}$$

- the best fit is rather curve not a straight line (attention: overfitting)



7.Polynominal Regression

- polynominal with higher degree can give us lower error rate
- if degree will be too high then overfitting will occur
- curve sholud fit the nature of the problem (trend) not every single sample

example frame

AGH

Thank you for your attention!

AGH

Q & A