MTAT.03.227 Machine Learning
Practice session 3
Basics of linear regression

September 23-25, 2019

1 Dataset

We continue to work with the animal dataset (cats and dogs are now together). Overall, today we will try to find the dependency (function) $Height(Weight)$.

	weight	height
1	4	2
2	5	1
3	5	2
4	5	3
5	6	2
6	7	4
7	11	8
8	11	10
9	13	8
10	13	10

Table 1: Animal data

2 Task 1: Errors

Calculate root mean squared error for the naive model $y = x - 3$, using one of the formulae below.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|. \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2. \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}. \tag{3}$$

Question: In what sense $RMSE$ is better than MSE?

3 Task 2: Regression

Calculate α and β for the regression line $height_i = \alpha + \beta \times weight_i + \varepsilon_i$ using the formulae below. Calculate RMSE for the model you have got. Is it lower than for the naive model?

$$y_i = \alpha + \beta x_i + \varepsilon_i. \tag{4}$$

$$\hat{\beta} = \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2} = \frac{\text{Cov}[x,y]}{\text{Var}[x]} \tag{5}$$

$$\hat{\alpha} = \overline{y} - \hat{\beta}\,\overline{x} \tag{6}$$

4 Task 3: Outlier

Add point $(10, 13)$ to the data. Calculate the new line equation based on 11 points in total. How this outlier affects regression?
Question: How can we regularize our regression to reduce outlier effect?

5 Homework: Multivariate notation

For multivariate task we will use the respective notation as follows. Linear regression has so called closed form solution (9), you will use it to calculate regression coefficients in your homework.

$$\sum_{j=1}^{n} X_{ij}\beta_j = y_i,\ (i = 1, 2, \ldots, m) \longrightarrow \mathbf{X}\beta = \mathbf{y} \tag{7}$$

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}. \tag{8}$$

$$\hat{\beta} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}. \tag{9}$$