

# INF 5860 Week 2

## Solution hints

Please note that the solution hints are in terms of keywords, and during an exam you might elaborate some more.

### Exercise 1 Linear regression

a) What is the loss function for linear regression?

The squared error  $J(\theta) = \frac{1}{2m} \sum_{i=1}^m (\phi_i - y_i)^2$

b) How does gradient descent update the estimate, give the general formulae?

$$\theta_j = \theta_j - \varepsilon \frac{\partial}{\partial \theta_j} J(\theta^0, \theta^1)$$

c) Given  $x = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ ,  $y = \begin{bmatrix} 1.5 \\ 2 \\ 2.5 \end{bmatrix}$   
Plot x,y as points in a plot.

d) If we start with  $\theta_0=0$  and  $\theta_1=0$ , compute the value of the initial loss function

$$J = 1/6(1.5*1.5 + 2*2 + 2.5*2.5) = 12.5/6 = 2.08$$

e) If we start with  $\theta_0=0$  and  $\theta_1=0$ , compute the estimate after one iteration if the learning rate is 1.

$$\begin{aligned} \frac{\partial}{\partial w} J(w, b) &= \frac{\partial}{\partial w} \frac{1}{2m} \sum_i (wx_i + b - y_i)^2 & \frac{\partial}{\partial b} J(w, b) &= \frac{\partial}{\partial b} \frac{1}{2m} \sum_i (wx_i + b - y_i)^2 \\ \frac{2}{2m} \sum_i (wx_i + b - y_i) x_i &= 1/3 * (1 * 1.5 + 2 * 2 + 2.5 * 3) = 13/3 = 4.33 & \frac{2}{2m} \sum_i (wx_i + b - y_i) &= 1/3 * (1.5 + 2 + 2.5) = 2 \end{aligned}$$

So with learning rate 1, and  $\theta_0=0$  and  $\theta_1=0$ , after 1 iteration the estimates are  $w=-4.33$  and  $b=-2$ .