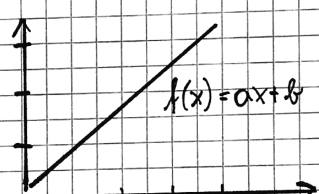
Exam Question 2:

Consider a set of observations with two features (X_1, X_2). A set of measurements for a class is composed of the observations $(2,-1)$ and $(3,-1)$. Another distinct set of measurements for a different class is composed of Observations $(-1,3)$ and $(-1,1)$.
We would like to use a multivariate linear regression to obtain a line that discriminate the two sets.

a) Explain in your own words what a multivariate linear regression is?

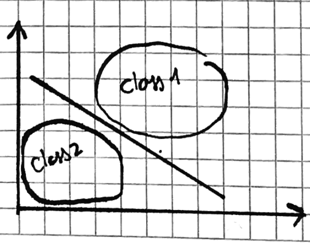
- In statistics, linear regression is a linear approach to modelling the relationship between dependent variables and one ore more independent variables.
- a linear regression model assumes that the relationship between the dependent variable y and the p-predictor is linear x.
  A simple linear function could look like this $f(x) = ax + b$



- linear regression is a supervised task an computes a continous output value.

b) How can we discriminate out of a regression a classification task?

- By using logistic regression. It's a classification algorithm that maps observations to a discrete set of classes: typically two.
  - linear regression computes a continous output value
  - logistic regression computes a probability



Logistic Regression

c) prepare your data to train a linear classifier in the form $y = W^T X$ using a stocastic gradient descent algorithm.

- We need to set up for all the terms in the equation $y = W^T X$. Lets start with y.
  - We can say that the set composed of observations $(2, -4)$ and $(3, -1)$ to be of class $y = -1$ (This has to be numeric) and the observations $(-1, 3)$ and $(-1, 1)$ of class $y = 1$.
  - First assign class -1 and 1 to the values (labeling)

We make the vector $W^T = [w_0, w_1, w_2]$ where w_0 stands for the bias and w_1, w_2 are the weights for features x_1, x_2. As there has been no initiard weights, we can start making this vector all zeros $W^T = [0, 0, 0]$

The vector for data X is composed of augmentations of each observation where we make each observation $[1, x_1, x_2]^T$ so that $y = w_0 + w_1 x_1 + w_2 x_2$

1) assign classes
2) assign weigts $= W^T = [0, 0, 0]$
3) Formula

$$g(x) = w^t x + w_0 \Rightarrow y = w_0 + w_1 + w_2 \begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix}$$

4) Make the matrices

$$X = \begin{bmatrix} 1 & 2 & -4 \\ 1 & 3, & -1 \\ 1 & -1 & 3 \\ 1 & -1 & 1 \end{bmatrix} \quad y = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

5) Formula ?!

$$\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = [0, 0, 0] \begin{bmatrix} 1, & 2 & -4 \\ 1, & 3 & -1 \\ 1, & -1 & 3 \\ 1, & 1 & 1 \end{bmatrix}$$

b) we need to visit each data from X, say X_i and make a prediction $y_p = WX_i$.

$e = y_p - y_i$ and then update $W_{i+1} = W_i - 0{,}1 e X_i$, where 0,1 is the learning factor.

1) Step 0 = initialize the weight

2) Step 1 = $w_T = [0, 0, 0] \times [1, 2, -4] = 0 - (-1) = 1$

$W_{i+1} = W_i - 0{,}1 \cdot 1 \cdot 1$

$W_{1+1} = 0 - 0{,}1 \times 1 \times 1 = -0{,}1$

$W_{2+1} = 0 - 0{,}1 \times 1 \times 2 = -0{,}2$

$W_{3+1} = 0 - 0{,}1 \times 1 \times (-4) = 0{,}4$

3) Step 2 = $W_i\ [-0{,}1, -0{,}2, 0{,}4] \times [1\ 3\ -1] = -1{,}1$

1) compute the error $-1{,}1 - (-1) = -0{,}1$

2) Update W as $W = W - 1 \times e \times X$

3)

$W_{1+1} = (-0{,}1) - 0{,}1 \times (-0{,}1) \times 1 = -0{,}09$

$W_{2+1} = (-0{,}2) - 0{,}1 \times (-0{,}1) \times 3 = -0{,}17$

$W_{3+1} = (0{,}4) - 0{,}1 \times (-0{,}1) \times (-1) = 0{,}39$

4) Step 3 $[-0{,}09, -0{,}17, 0{,}39] \times [1\ -1\ 3] = 1{,}25$

$1{,}25 - (1) = 0{,}25$

$W_{1+1} = (-0{,}09) - 0{,}1 \times (0{,}25) \times 1 = -0{,}115$

$W_{2+1} = (-0{,}17) - 0{,}1 \times (0{,}25) \times -1 = -0{,}14$

$W_{3+1} = (0{,}39) - 0{,}1 \times (0{,}25) \times 3 = 0{,}31$

5) step 1) $[-0{,}11, -0{,}14, 0{,}31] \times [1, -1, 1] = 0{,}34 - 1 = -0{,}65$

$w_1+1 = (-0{,}11) - 0{,}1 \times (-0{,}65) \times 1 = -0{,}04$

$w_2+1 \quad (-0{,}14) - 0{,}1 \times (-0{,}65) \times -1 = -0{,}21$

$w_0+1 \quad (0{,}31) - 0{,}1 \times (-0{,}65) \times 1 = 0{,}38$

7) New weight after one epoch = $W = [-0{,}04, -0{,}21, 0{,}38]$

Make a prediction, after one epoch of training, as to which class observation (1, -1) belongs to.

1) How does now our equation look like?

$$y_p = -0{,}04 + (-0{,}21) + (0{,}38) \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

or

$$y = [-0{,}04 \;\; -0{,}21 \;\; 0{,}38][1 \;\; 1 \;\; -1]^T =$$

or

$$y = (-0{,}04 \times 1 + (-0{,}21 \times 1) + (0{,}38 \times -1) = -0{,}66$$

belongs to class -1