Machine Learning & Machine Learning (extended)

Practice Exercise Sheet - Linear Regression/Modelling

Question (Book exercise 1.2): Write a MATLAB script that can find linear model parameters w_0 and w_1 for an arbitrary dataset of x_n, t_n pairs. Use this script to learn a model for Olympics 100m men's data.

Question (Book exercise 1.3): Show that:

$$\mathbf{w}^{\mathsf{T}}\mathbf{X}^{\mathsf{T}}\mathbf{X}\mathbf{w} = w_0^2\left(\sum_{n=1}^{N} x_{n1}^2\right) + 2w_0w_1\left(\sum_{n=1}^{N} x_{n1}x_{n2}\right) + w_1^2\left(\sum_{n=1}^{N} x_{n2}^2\right),$$

where

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, \ \mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{bmatrix}.$$

Hint: it's probably easiest to do the $\boldsymbol{X}^T\boldsymbol{X}$ first!

Question (Book exercise 1.6): Write a MATLAB script that can find linear model parameters $\boldsymbol{w}$ for an arbitrary dataset of $\boldsymbol{x}_n$, t_n pairs with $\boldsymbol{x}_n$ being of arbitrary dimensions. Use this script to learn a model for Olympics 100m men's data.

Question: Suppose we have a set of values as below. Is it possible to find a function f such that $y = f(x)$? Can you describe this function f and suggest how it can be estimated from the data given?

x	y
1.2	1.1
2.3	2.1
3.0	3.1
3.8	4.0
4.7	4.9
5.9	5.9

Question: What is model complexity in linear modelling? If the model complexity increases, does it help or hinder function modelling? Comment on both scenarios.

Question (Book exercises 1.7 and 1.8): Given the linear models below to predict winning time t from Olympic year x for 100m race:

Men's model: $t = 36.4165 - 0.0133x$

Women's model: $t = 40.9242 - 0.0151x$

a) Can you predict what year the Women's winning time will be lower than the Men's winning time?

b) What are the predicted winning times for both Men's race and Women's race for that year? Comment on that.

Question: Let's consider that a linear model (as a line) has two parameters $\widehat{w}_0$ and $\widehat{w}_1$ which can be estimated as below:

$$\widehat{w}_0 = \bar{t} - \widehat{w}_1\bar{x}$$

$$\widehat{w}_1 = \frac{\overline{xt} - \bar{x}\bar{t}}{\overline{x^2} - (\bar{x})^2}$$

a) Given the data below, can you estimate the model parameters using above equations.

x	t
1.2	1.1
2.3	2.1
3.0	3.1
3.8	4.0
4.7	4.9
5.9	5.9

b) Having found the model parameters, can you use the model to predict target labels t?

c) Is there a way to find how good the model is doing by comparing model predictions with actual targets?

Question (Book exercise 1.10): Given the total training loss estimate function below between actual target labels t_n and model predicted target labels $\boldsymbol{w}^T\boldsymbol{x}_n$ where $\boldsymbol{x}_n$ denote the attributes and $\boldsymbol{w}$ denote model parameters:

$$\mathcal{L} = \sum_{n=1}^{N}(t_n - \boldsymbol{w}^T\boldsymbol{x}_n)^2 = (\boldsymbol{t} - \boldsymbol{X}\boldsymbol{w})^T(\boldsymbol{t} - \boldsymbol{X}\boldsymbol{w})$$

Can you mathematically derive the expression for parameters estimate $\widehat{\boldsymbol{w}}$ for linear modelling that minimizes this total loss? How does it compare with the parameters $\widehat{\boldsymbol{w}}$ for linear model that minimizes average loss?

Question: A linear model can attempt to consider the noise present in the following way:

$t_n = \boldsymbol{w}^T\boldsymbol{x}_n + \varepsilon_n$ where ε_n denotes the additive noise.

a) Why is the noise considered to be additive?

b) Is there a way to model the statistical distribution of this noise?