COMPSCI 4ML3 - Introduction to Machine Learning
Assignment 1 Solution

October 14, 2024

Alireza Daeijavad
McMaster University

---

1. [**70 points**] Programming Component. Follow this link to open the Google colab environment and make a copy of the notebook. Include the answers/graphs/pictures/analyses the tasks in your final pdf report. Additionally, upload your modified Jupyter notebook that includes your code (as a separate ipynb file).

> **Solution:** LINK

2. [**15 points**] In this question we will use least squares to find the best line ($\hat{y} = ax + b$) that fits a non-linear function, namely $f(x) = 2x - x^3 - 1$. For this, assume that you are given a set of n training points $\{(x^i, y^i)\}_{i=1}^n$=$\{((i/n), 2(i/n) - (i/n)^3) - 1\}_{i=1}^n$. Find a line (i.e., $a, b \in \mathbb{R}$) that fits the training data the best when $n \to \infty$. Write down your calculations as well as the final values for a and b. (Additional notes: the $n \to \infty$ assumption basically means that we are dealing with an integral rather than a finite summation. If it makes it easier for you, instead of working with actual training data you can assume x is uniformly distributed on $[0, 1]$.)

> **Solution:** Let $f(x)$ denote *pdf* of x. Since x is uniformly distributed between $[0, 1]$, $f(x) = 1/(1 - 0) = 1$.
>
> $$\begin{aligned} \min_{a,b} g &= \min_{a,b} \int_0^1 (ax + b - y)^2 f(x)dx \\ &= \min_{a,b} \int_0^1 (ax + b - 2x + x^3 + 1)^2 dx \end{aligned}$$
>
> When $\frac{\partial g}{\partial a} = 0$, we have
>
> $$\begin{aligned} \int_0^1 2x(ax + b - 2x + x^3 + 1)dx &= 0 \\ (2/3ax^3 + bx^2 - 4/3x^3 + 2/5x^5 + x^2)\Big|_0^1 &= 0 \\ 2/3a + b - 4/3 + 2/5 + 1 &= 0 \\ 10a + 15b - 20 + 6 + 15 &= 0 \\ 10a + 15b &= -1 \end{aligned}$$

When $\frac{\partial g}{\partial b} = 0$, we have

$$\begin{aligned}
\int_0^1 2(ax + b - 2x + x^3 + 1)dx &= 0 \\
ax^2 + 2bx - 2x^2 + 1/2x^4 + 2x\Big|_0^1 &= 0 \\
a + 2b - 2 + 1/2 + 2 &= 0 \\
2a + 4b - 4 + 1 + 4 &= 0 \\
2a + 4b &= -1
\end{aligned}$$

Therefore, $a = 1.1$, and $b = -0.8$.

3. [**15 points**] In this question we would like to fit a line with zero y-intercept ($\hat{y} = ax$) to the curve $y = x^2$. However, instead of minimizing the sum of squares of errors, we want to minimize the following objective function:

$$\sum_i \left[\log\left(\frac{\hat{y}^i}{y^i}\right)\right]^2$$

Assume that the distribution of x is uniform on $[2, 4]$. What is the optimal value for a? Show your work.

Solution: Let $f(x)$ denote *pdf* of x. Since x is uniformly distributed between $[2, 4]$, $f(x) = 1/(4-2) = 0.5$.

$$\begin{aligned}
\min_a g &= \min_a \int_2^4 (\ln(\frac{ax}{x^2}))^2 f(x)dx \\
&= \min_a \int_2^4 0.5(\ln(\frac{a}{x}))^2 dx \\
&= \min_a \int_2^4 0.5(\ln(a) - \ln(x))^2 dx
\end{aligned}$$

When $\frac{\partial g}{\partial a} = 0$, we have

$$\begin{aligned}
\int_2^4 \frac{\ln(a)}{a} - \frac{\ln(x)}{a} dx &= 0 \\
\frac{\ln(a)}{a}x - \frac{1}{a}\left(x\ln(x) - x\right)\Big|_2^4 &= 0 \\
4\ln(a) - (4\ln(4) - 4) - 2\ln(a) + (2\ln(2) - 2) &= 0 \\
\ln(a) - 3\ln(2) + \ln(e) &= 0 \\
\ln(\frac{ae}{8}) &= 0 \\
a &= \frac{8}{e}
\end{aligned}$$

Therefore, $a = \frac{8}{e} = 2.943$.