MAD - assignment 2

Frederik Altmann

November 2024

1 Exercise 1

1.1 a)

$$\begin{aligned}
\mathcal{L} &= \frac{1}{N}\sum_{n=1}^{N}\alpha_n(f(x_n;w)-t_n)^2 \\
&= \frac{1}{N}||A||Xw-t||^2||
\end{aligned}$$

Here:

- X the matrix of stacked inputs x_n of size Nxd
- w is a d-dimensional vector
- t is a N-dimensional vector
- A is a diagonal matrix that contains weights $\alpha_1, \cdots, \alpha_N$ on the diagonal line

$$\begin{aligned}
\mathcal{L} &= \frac{1}{N}A(Xw-t)(Xw-t)^T) \\
\frac{\partial\mathcal{L}}{\partial w} &= \frac{2}{N}A(Xw-t)X^T \\
&\text{Set gradient to 0} \\
0 &= \frac{2}{N}A(Xw-t)X^T \\
0 &= A(Xw-t)X^T \\
0 &= AXX^Tw - AtX^T \\
AtX^T &= AXX^Tw \\
(AXX^T)^{-1}AtX^T &= w
\end{aligned}$$

1.2 b

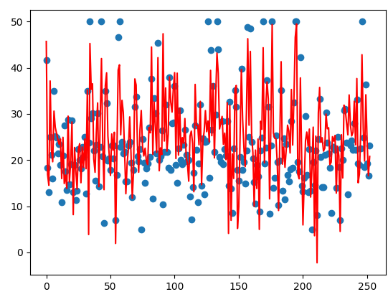
To do this I replaced, these 2 lines in the fit function:

```
A = numpy.diag(t.reshape(-1)**2)
self.w = numpy.linalg.inv((X.T @ A) @ X) @ (X.T @ A) @ t
```

And return this in the predict step:

```
X @ self.w
```

When plotting these, I get:



What do you expect to happen?

I expect that the additional weights $\alpha_n = t_n^2$ give more importance to larger target values, which means it will fits closer the these points.

What do you observe?

I can see in the above plot that the model predicts much better on values with higher target values. This may be due to the higher weights.

Do the additional weights have an influence on the outcome?

Yes, as mentioned before, the model predicts much better values with higher target values. This is because of the higher weights.

2 Exercise 2

2.1 a)

The best value of $\lambda = 7.564 \cdot 10^{-5}$, which gives loss 0.0521

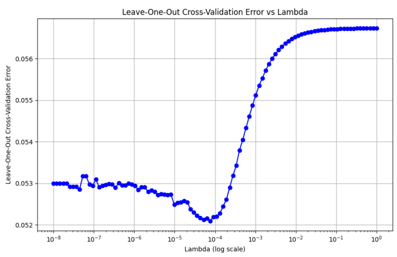
Coefficients

1. **Without regularization** ($\lambda = 0$):

$$[23681.81, 36.14, 0.0184, 3.122 \cdot 10^6]$$

2. **With Best Regularization** ($\lambda = 7.564 \cdot 10^{-5}$)

$$[1.557 \cdot 10^4, -23.6725, 0.012, -2.0296 \cdot 10^{-06}]$$



2.2 b)

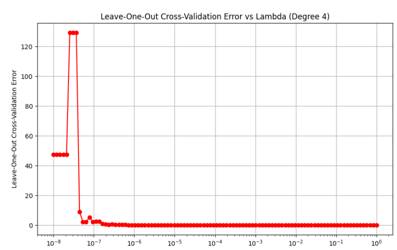
The best value of $\lambda = 4.641 \cdot 10^{-6}$, which gives loss 0.0468
Coefficients

1. **Without regularization** ($\lambda = 0$):

$$[2.3681 \cdot 10^4, -3.6143 \cdot 10^1, 1.839 \cdot 10^2, -3.1221 \cdot 10^6]$$

2. **With Best Regularization** ($\lambda = 4.641 \cdot 10^{-6}$)

$$[1.0231 \cdot 10^4, -8.511, -2.888 \cdot 10^{-3}, 4.1649 \cdot 10^{-6}, -9.3542 \cdot 10^{-10}]$$



3 Exercise 3

3.1 a)

1. **Diff** $F(x)$ for $x > 0$

$$f(x) = \frac{dF(x)}{dx} = \frac{d(1 - exp(-\beta x^\alpha))}{dx}$$

2. **Apply the chain rule**:

$$f(x) = \beta\alpha x^{\alpha-1} exp(-\beta x^\alpha, x > 0$$

3. So the pdf is:

$$f(x) = \begin{cases} 0 & x \leq 0, \\ \beta\alpha x^{\alpha-1} e^{-\beta x^\alpha} & x > 0. \end{cases}$$

3.2 b)

3.2.1 i

$$
\begin{aligned}
P(X > 4) &= 1 - F(4) \\
\text{Compute } x = 4 \\
F(4) &= 1 - exp(-\frac{1}{2} \cdot 4^2) \\
&= 1 - exp(8) \\
P(X > 4) &= exp(-8)
\end{aligned}
$$

3.2.2 ii

$$
\begin{aligned}
P(5 \leq X \leq 10 &= F(10) - F(5) \\
\text{Compute } F(10) \text{ and } F(5) \\
F(10) &= 1 - exp\left(-\frac{1}{2} \cdot 10^2\right) \\
&= 1 - exp(-50) \\
F(5) &= 1 - exp\left(-\frac{1}{2} \cdot 5^2\right) \\
&= 1 - exp(-12.5) \\
P(5 \leq X \leq 10 &= (1 - exp(-50)) - (1 - exp(-12.5)) \\
&= exp(-12.5) - exp(-50)
\end{aligned}
$$

3.3 c)

The median is the value x_m such that $F(x_m) = 0.5$

$$
\begin{aligned}
1 - exp(-\beta x_m^\alpha) &= 0.5 \\
exp(-\beta x_m^\alpha) &= 0.5 \\
\text{Take the natural logarithm} \\
\beta x_m^\alpha &= ln(0.5) \\
x_m^\alpha &= -\frac{ln(0.5)}{\beta} \\
x_m &= (-\frac{ln(0.5)}{\beta})^{\frac{1}{\alpha}}
\end{aligned}
$$

4 Exercise 4

Part (a) 1. Remaining Silent:
The probability of conviction is:

$$P(\text{Conviction — Silent, NC}) = P(\text{Court — Silent, NC}) \cdot \big(1 - P(\text{Acquittal — Silent, NC})\big)$$

Substituting the values:

$$\begin{aligned} P(\text{Conviction — Silent, NC}) &= 0.001 \cdot (1 - 0.8) \\ &= 0.001 \cdot 0.2 \\ &= 0.0002 \end{aligned}$$

The expected prison time is:

$$\begin{aligned} E(\text{Prison Days — Silent, NC}) &= P(\text{Conviction — Silent, NC}) \cdot 1825 \\ E(\text{Prison Days — Silent, NC}) &= 0.0002 \cdot 1825 = 0.365\,\text{days} \end{aligned}$$

2. Talking to the Police:
The probability of conviction is:

$$P(\text{Conviction — Talk, NC}) = P(\text{Court — Talk, NC}) \cdot \big(1 - P(\text{Acquittal — Talk, NC})\big)$$

Substituting the values:

$$\begin{aligned} P(\text{Conviction — Talk, NC}) &= 0.0015 \cdot (1 - 0.2) \\ &= 0.0015 \cdot 0.8 \\ &= 0.0012 \end{aligned}$$

The expected prison time is reduced by 50%, resulting in:

$$\begin{aligned} E(\text{Prison Days — Talk, NC}) &= P(\text{Conviction — Talk, NC}) \cdot 912.5 \\ E(\text{Prison Days — Talk, NC}) &= 0.0012 \cdot 912.5 \\ &= 1.095\,\text{days} \end{aligned}$$

Conclusion for Part (a):

- Remaining Silent: 0.365 days.
- Talking to the Police: 1.095 days.

It is better to remain silent.

Part (b) 1. Remaining Silent:

The probability of conviction is:

$$P(\text{Conviction — Silent, C}) = P(\text{Court — Silent, C}) \cdot \big(1 - P(\text{Acquittal — Silent, C})\big)$$

Substituting the values:

$$P(\text{Conviction — Silent, C}) = 0.005 \cdot (1 - 0.2) = 0.005 \cdot 0.8 = 0.004$$

The expected prison time is:

$$E(\text{Prison Days — Silent, C}) = P(\text{Conviction — Silent, C}) \cdot 1825$$

$$E(\text{Prison Days — Silent, C}) = 0.004 \cdot 1825 = 7.3\,\text{days}$$

2. Talking to the Police: The probability of conviction is:

$$P(\text{Conviction — Talk, C}) = P(\text{Court — Talk, C}) \cdot \big(1 - P(\text{Acquittal — Talk, C})\big)$$

Substituting the values:

$$P(\text{Conviction — Talk, C}) = 0.005 \cdot (1 - 0.05) = 0.005 \cdot 0.95 = 0.00475$$

The expected prison time is reduced by 50%, resulting in:

$$E(\text{Prison Days — Talk, C}) = P(\text{Conviction — Talk, C}) \cdot 912.5$$

$$E(\text{Prison Days — Talk, C}) = 0.00475 \cdot 912.5 = 4.335\,\text{days}$$

Conclusion for Part (b):

- Remaining Silent: $7.3\,\text{days}$.
- Talking to the Police: $4.335\,\text{days}$.

It is better to talk to the police.

Final Summary:

- Part (a): For a person with no history of convictions (NC), it is better to **remain silent**.
- Part (b): For a person with a history of convictions (C), it is better to **talk to the police**.