MAD 2022-23, Assignment 1

Helga Rykov Ibsen <mcv462> Hold 6

1. september 2023

Opgave 1

(a)

Vi er givet funktionen:

$$f(x, y) = x^4y^3 + x^5 - e^y$$

Vi finder de partielle afledede af f mht. x og y som:

$$\frac{\partial f}{\partial x} = 4x^3y^3 + 5x^4$$
$$\frac{\partial f}{\partial y} = 3x^4y^2 - e^y$$

(b)

Vi er givet funktionen:

$$f(x, y) = \frac{1}{\sqrt{x^3 + x \cdot y + y^2}}$$

Vi finder de partielle afledede af f mht. x ved at bruge kedereglen som:

$$\begin{aligned}\frac{\partial f}{\partial x} &= -\frac{\frac{1}{2\sqrt{x^3+x\cdot y+y^2}} \cdot (3x^2 + y)}{(\sqrt{x^3 + x \cdot y + y^2})^2} \\ &= -\frac{3x^2 + y}{2\sqrt{x^3 + x \cdot y + y^2} \cdot (x^3 + x \cdot y + y^2)}\end{aligned}$$

Vi finder de partielle afledede af f mht. y ved at bruge kedereglen som:

$$\begin{aligned}\frac{\partial f}{\partial y} &= -\frac{\frac{1}{2\sqrt{(x^3+x\cdot y+y^2)}} \cdot (x + 2y)}{(\sqrt{x^3 + x \cdot y + y^2})^2} \\ &= -\frac{x + 2y}{2\sqrt{x^3 + x \cdot y + y^2} \cdot (x^3 + x \cdot y + y^2)}\end{aligned}$$

(c)

Vi er givet funktionen:

$$f(x,y) = \frac{x^3+y^2}{x+y}$$

Vi finder de partielle afledede af f mht. x ved at bruge brøkreglen for differentiation som:

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{(x+y)\cdot(x^3+y^2)' - (x^3+y^2)\cdot(x+y)'}{(x+y)^2} \\ &= \frac{3x^2(x+y) - (x^3+y^2)}{(x+y)^2}\end{aligned}$$

Vi finder de partielle afledede af f mht. y som:

$$\begin{aligned}\frac{\partial f}{\partial y} &= \frac{(x+y)\cdot(x^3+y^2)' - (x^3+y^2)\cdot(x+y)'}{(x+y)^2} \\ &= \frac{2y(x+y) - (x^3+y^2)}{(x+y)^2}\end{aligned}$$

Opgave 2

(a)

Vi er givet funktionen:

$$f(\overline{x}) = \overline{x}^T \cdot \overline{x} + c$$

Vi anvender tabellen (s.23 i bogen) og finder gradienten for f mht. $\overline{x}$ som følger:

$$\frac{\partial f}{\partial \overline{x}} = 2\overline{x}$$

(b)

Vi er givet funktionen:

$$f(\overline{x}) = \overline{x}^T \cdot \overline{b}$$

Vi anvender tabellen (s.23 i bogen) og finder gradienten for f mht. $\overline{x}$ som følger:

$$\frac{\partial f}{\partial \overline{x}} = \overline{b}$$

(c)

Vi er givet funktionen:

$$f(\overline{x}) = \overline{x}^T \cdot A\overline{x} + \overline{b}^T \cdot \overline{x} + c$$

Vi anvender tabellen (s.23 i bogen) og finder gradienten for f mht. $\overline{x}$ som følger:

$$\frac{\partial f}{\partial \overline{x}} = 2A\overline{x} + \overline{b}$$

Opgave 3

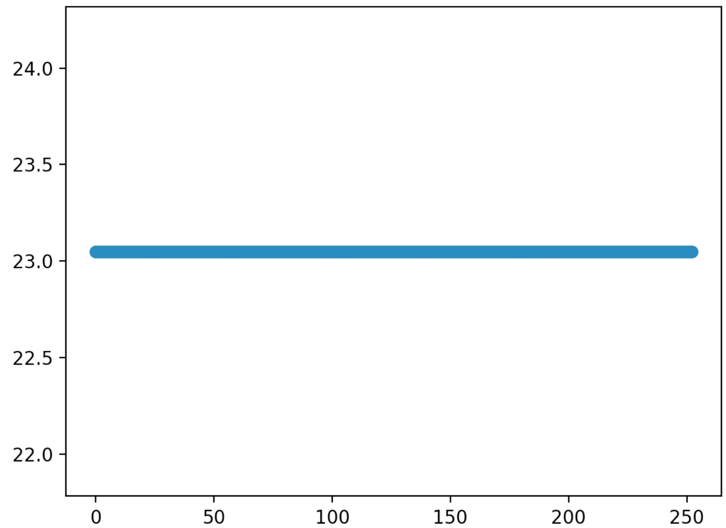
(a)

Figur 1: The mean of the house prices

(b)

Figur 2: The value of the RMSE

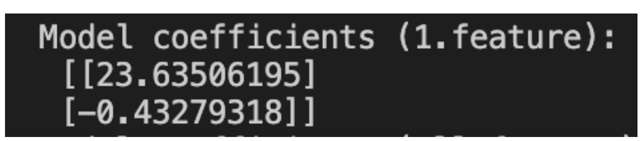
(c)



Figur 3: The 2D scatter plot

Opgave 4

(b)



```
Model coefficients (1.feature):
 [[23.63506195]
 [-0.43279318]]
```

Figur 4: Weights w_0 and w_1

De to koefficienter fortæller os om hvordan modellen/udviklingen ser ud grafisk. Den første konstante koefficient w_0 fortæller os om hvor den rette linje kommer til at ligge ift.y-aksen (ca. 23) og den anden konstante koefficient w_1 fortæller os at den rette linje kommer til at gå i faldende/nedadgående retning fordi den er negativ.

(c)

```
Model coefficients (all feature):
 [[ 1.95123661e+01]
 [-6.65962795e-02]
 [ 1.65961371e-02]
 [-5.18606201e-02]
 [ 2.82306875e+00]
 [-2.14548721e+01]
 [ 6.09191498e+00]
 [-4.38038546e-02]
 [-1.42409460e+00]
 [ 2.87670030e-01]
 [-1.39327428e-02]
 [-9.33505242e-01]
 [ 1.54948728e-02]]
```

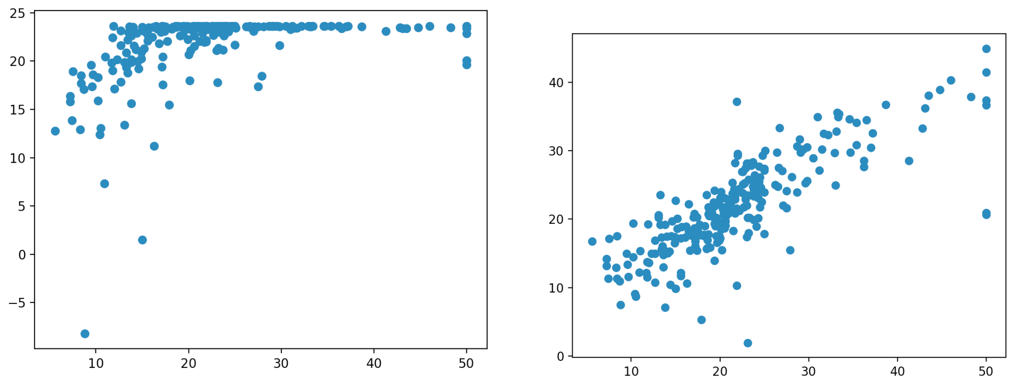
Figur 5: Weights w_i

(d)

```
The value of the loss function for the first feature: 7

The value of the loss function for all features: 5
```

Figur 6: The value of the RMSE



(a)

(b)

Figur 7: The 2D scatter plots of the model based on the first feature (a) and all features (b)

Opgave 5

Vi skal finde de optimale estimater $\hat{w}_0$ og $\hat{w}_1$ for den *totale training* funktion $\mathcal{L}$:

$$\mathcal{L} = \sum_{n=1}^{N} (f(\mathbf{x}_n; \mathbf{w}) - t_n)^2 = \sum_{n=1}^{N} (\mathbf{w}^T \mathbf{x}_n - t_n)^2$$

Vi starter med at omskrive det første led i parentesen i det sidste udtryk på formen af ligningen for den rette linje:

$$\begin{aligned}
\mathcal{L} &= \sum_{n=1}^{N} (\mathbf{w}^T \mathbf{x}_n - t_n)^2 \\
&= \sum_{n=1}^{N} (w_0 + w_1 x_n - t_n)^2 \\
&= \sum_{n=1}^{N} (w_0^2 + w_1^2 x_n^2 + t_n^2 + 2w_0 w_1 x_n - 2w_1 t_n x_n - 2w_0 t_n)
\end{aligned}$$

Vi skal nu differentiere udtrykket mht. variablerne $\hat{w}_0$ og $\hat{w}_1$:

$$\frac{\partial \mathcal{L}}{\partial w_0} = \sum_{n=1}^{N} (2w_0 + 2w_1 x_n - 2t_n)$$

$$\frac{\partial \mathcal{L}}{\partial w_1} = \sum_{n=1}^{N}(2w_1 x_n^2 + 2w_0 x_n - 2t_n x_n)$$

Nu skal vi løse ligningssystemet for $\hat{w}_0$ og $\hat{w}_1$. Vi sætter de to ligninger lige med 0, smider det sidste led på den anden side af lighedstegnet og dividerer med 2 på begge sider:

$$\sum_{n=1}^{N} w_0 + \sum_{n=1}^{N} w_1 x_n = \sum_{n=1}^{N} t_n$$

$$\sum_{n=1}^{N} w_1 x_n^2 + \sum_{n=1}^{N} w_0 x_n = \sum_{n=1}^{N} t_n x_n$$

Da summen af elementerne x er det sammen som at gange antallet af elementer N med gennemsnittet $\overline{x}$ kan vi omskrive de to ligninger (n er gemt i gennemsnittet og kan også bortforkortes) som :

$$N w_0 + w_1 N \overline{x} = N \overline{t}$$

$$w_1 N \overline{x^2} + w_0 N \overline{x} = N \overline{tx}$$

Da N optræder alle steder kan den også bortforkortes:

$$w_0 + w_1 \overline{x} = \overline{t}$$

$$w_1 \overline{x^2} + w_0 \overline{x} = \overline{tx}$$

Nu starter med at isolere w_0 i den første ligning og bestemme udtrykket for det første optimale estimatet $\hat{w}_0$:

$$\hat{w}_0 = \overline{t} - w_1 \overline{x}$$

For at finde estimatet $\hat{w}_1$ indsætter vi det fundne $\hat{w}_0$:

$$w_1 \overline{x^2} + (\overline{t} - w_1 \overline{x})\overline{x} = \overline{tx}$$

$$w_1 \overline{x^2} + \overline{t}\overline{x} - w_1 \overline{x}^2 = \overline{tx}$$

$$w_1(\overline{x^2} - \overline{x}^2) = \overline{tx} - \overline{t}\overline{x}$$

$$w_1 = \frac{\overline{tx} - \overline{t}\overline{x}}{\overline{x^2} - \overline{x}^2}$$

Vi kan hermed se at de fundne estimater for den totale *loss training* funktion er de samme som for den gennemsnitlige en (se bogen s.26). Det forventede vi også, fordi når vi skal finde parametrene som optimerer *loss*, så er det lige meget om vi bruger gennemsnittet af $x'erne$ og $t'erne$ eller totalt antal af dem.