{Q4 Gradients}

* • The rest are answered in the notebook (Q1, Q2, Q3)

Define $f: \mathbb{R}^d \longrightarrow \mathbb{R}$, $f(\omega) = \omega^T x + b$

for a given $x \in \mathbb{R}^d$ and $b \in \mathbb{R}$

① $\nabla_\omega f = \left(\frac{df}{dw_1}, \ldots, \frac{df}{dw_d}\right)^T = (x_1, \ldots, x_d)^T = x$ ■

② The Hessian matrix of f is:

$$H_f = \begin{pmatrix} \frac{d^2 f}{dw_1^2} & \frac{d^2 f}{dw_1 dw_2} & \cdots & \frac{d^2 f}{dw_1 dw_n} \\ \vdots & & & \\ \frac{d^2 f}{dw_n dw_1} & \cdots & \cdots & \frac{d^2 f}{dw_n dw_n} \end{pmatrix}$$

$= 0_{n \times n}$ = Zero matrix

③ The zero matrix is PSD

To be continued

Q4

(4) define $g: \mathbb{R}^d \to \mathbb{R}$, $\lambda > 0$,

$$g(w) = \frac{1}{2}\lambda \|w\|^2$$

$$\nabla_w g = \nabla_w\left(\frac{\lambda}{2} w^T w\right) = \frac{\lambda}{2}\left(\nabla_w\left(\sum_{i=1}^{d} w_i^2\right)\right)$$

$$= \frac{\lambda}{2}\left(2w_1, \dots, 2w_n\right) = \lambda \cdot w$$

(5)
$$H_g = \begin{pmatrix} \frac{d^2 g}{dw_1^2} & \cdots & \frac{d^2 g}{dw_1 dw_n} \\ \vdots & & \\ \frac{d^2 g}{dw_n w_1} & \cdots & \frac{d^2 g}{dw_n^2} \end{pmatrix} = \lambda Id_{d \times d}$$

$$\frac{d^2 g}{dw_i dw_j} = \begin{cases} \lambda & \longleftarrow, i = j \\ 0 & \longleftarrow, i \neq j \end{cases}$$

(6) $\lambda Id_{d\times d}$ is PD, since $\lambda > 0$. ∎

Scanned with CamScanner

Q4 **7**

$$h(w_1, w_2) = 12 w_1^3 - 36 w_1 w_2 - 2 w_2^3 + 9 w_2^2 - 72 w_1 + 60 w_2 + 5$$

$\Rightarrow$ I $\quad \dfrac{dh}{dw_1} = 36 w_1^2 - 36 w_2 - 72 = 0$

needed

II $\quad \dfrac{dh}{dw_2} = -36 w_1 - 6 w_2^2 + 18 w_2 + 60 = 0$

From I, $w_2 = w_1^2 - 2$, plugging that into II, we get the constraint:

$$-36 w_1 - 6(w_1^2 - 2)^2 + 18 w_2 + 60 = 0$$

More elegantly: $-36 w_1 - 6 w_1^4 + 42 w_1^2 = 0$

$\Rightarrow 6 w_1^4 - 42 w_1^2 + 36 w_1 = 0 \Rightarrow w_1 (6 w_1^3 - 42 w_1 + 36) = 0$

$\Rightarrow w_1 (w_1 - 1)(w_1 - 2)(w_1 + 3) = 0$ $\quad$ $\boxed{\text{To be continued}}$

Q4: ⑦ Therefore the critical points can be $(0,-2), (1,-1), (2,2), (-3,7)$

---

⑧ $\frac{dh^2}{dw_1^2} = 72w_1$, $\frac{dh^2}{dw_1 dw_2} = \frac{dh^2}{dw_2 dw_1} = -36$

$\frac{dh^2}{dw_2^2} = -12w_2 + 18$

Note: (h) meets the second partial derivative test. Since h is and it's derivatives up to second order are elemetary therefore continuous.

Using:

$D(x,y) = det(H(x,y))$

we get:

Saddles

$D((0,-2)) = 0 \cdot (12+18) - (-36)^2 < 0$

$D((1,-1)) = 72 \cdot 1 \cdot (-12(-1)+18) - (-36)^2 > 0$

$D((2,2)) = 72 \cdot 2 (-12 \cdot 2 + 18) - (-36)^2 < 0$

Next page

$D((-3,7)) = 72 \cdot (-3)(-12 \cdot 7 + 18) - (-36)^2 > 0$

Scanned with CamScanner

Q4: (8): In the case of the points $(1,-1)$, $(-3,7)$

Checking $\frac{dh^2}{dw_1^2}(w_1,w_2)$:

$\frac{dh^2}{dw_1^2}(1,-1) = 72 > 0 \Rightarrow$ $\boxed{(1,-1) \text{ is a local min}}$

$\frac{dh^2}{dw_1^2}(-3,7) = -3\cdot 72 < 0 \Rightarrow$ $\boxed{(-3,7) \text{ is a local max}}$

(9) Does h have global max or min?

Answer: Max: No   Min: No

Next page

Proof: Assuming we have a global max, points of the form $(w_1,0)$ give us $h(w_1,0) = 12w_1^3 - 72w_1 + 5$, which isn't bound

9: Min: ~~Assu~~ Observe points of the form $(0, w_2)$,

$$h(0, w_2) = -2w_2^3 + 9w_2^2 + 60w_2$$

which is not bound (for large enough w_2 it can go to $-\infty$)