

{Q 4 Gradients}



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

Define $f: \mathbb{R}^d \rightarrow \mathbb{R}$, $f(w) = w^T x + b$
for a given $x \in \mathbb{R}^d$ and $b \in \mathbb{R}$

① $\nabla_w f = \left(\frac{df}{dw_1}, \dots, \frac{df}{dw_d} \right)^T = (x_1, \dots, x_d)^T = x$

② The Hessian matrix of f is:

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

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

③ The zero matrix is PSD

To be continued

Q 4

4 define $g: \mathbb{R}^d \rightarrow \mathbb{R}$, $\lambda > 0$,
 $g(w) = \frac{1}{2} \lambda \|w\|^2$

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

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

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

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

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



$$h(\omega_1, \omega_2) = 12\omega_1^3 - 36\omega_1\omega_2 - 2\omega_2^3 + 9\omega_2^2 - 72\omega_1 + 60\omega_2 + 5$$

$$\Rightarrow \textcircled{\text{I}} \frac{dh}{d\omega_1} = 36\omega_1^2 - 36\omega_2 - 72 = 0$$

$$\textcircled{\text{II}} \frac{dh}{d\omega_2} = -36\omega_1 - 6\omega_2^2 + 18\omega_2 + 60 = 0$$

From Ⓘ $\omega_2 = \omega_1^2 - 2$, plugging that into

Ⓣ, we get the constraint:

$$-36\omega_1 - 6(\omega_1^2 - 2)^2 + 18\omega_2 + 60 = 0$$

More elegantly: $-36\omega_1 - 6\omega_1^4 + 42\omega_1^2 = 0$

$$\Rightarrow 6\omega_1^4 - 42\omega_1^2 + 36\omega_1 = 0 \Rightarrow \omega_1(6\omega_1^3 - 42\omega_1 + 36) = 0$$

$$\Rightarrow \omega_1(\omega_1 - 1)(\omega_1 - 2)(\omega_1 + 3) = 0$$

To be continued

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

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

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

Using:

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

we get:

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

Saddles

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

Next page

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

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

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

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

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

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

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: ~~Assa~~ Observe: points of
the form $(0, w_2)$,

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

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