IN3310 Week 4 Solution

Alex Binder

1 Coding

See the attached file :)

2 Disastrous Derivatives

$$f(X) = Xa \implies Df(X)[H] = Ha \in \mathbb{R}^{(d,1)}$$

$$f(X) = XX^\top \implies Df(X)[H] = HX^\top + XH^\top \in \mathbb{R}^{(d,d)}$$

$$f(X) = XCX \implies Df(X)[H] = HCX + XCH \in \mathbb{R}^{(d,d)}$$

$$f(X) = CXBX^\top AX \implies Df(X)[H] = CHBX^\top AX + CXBH^\top AX + CXBX^\top AH \in \mathbb{R}^{(d,d)}$$

$$\begin{aligned}
f(X) &= \begin{pmatrix} 1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & x_2^3 \\ \sin x_2 & x_1 \end{pmatrix} \\
&= (1 + x_2 \sin x_2,\ x_2^3 + x_1 x_2) \\
\implies Df(X)[H] &= \frac{\delta f}{\delta x_1} h_1 + \frac{\delta f}{\delta x_2} h_2 \\
&= (0, x_2) h_1 + (\sin x_2 + x_2 \cos(x_2), 3x_2^2 + x_1) h_2 \\
&= ((\sin x_2 + x_2 \cos(x_2)) h_2, x_2 h1 + (3x_2^2 + x_1) h_2)
\end{aligned}$$

3 3 dims

The vector $x_2 - x_1$ is a good choice to define a hyperplane separating x_1 and x_2. We can now look for w by assuming that $w = s(x_2 - x_1)$ for some scaling factor $s \in \mathbb{R}$. We want $f(x_1) = 1$ and $f(x_2) = -1$, this gives us the system

$$s(x_2 - x_1)\cdot x_1 + b = 1$$
$$s(x_2 - x_1)\cdot x_2 + b = -1$$

so that

$$s = \frac{-2}{\|x_2 - x_1\|^2}.$$

Similarly for the bias b

$$w\cdot x_1 = 1 - b$$
$$w\cdot x_2 = -1 - b,$$

from which we can deduce

$$b = -\frac{1}{2}w\cdot (x_1 + x_2).$$

Plugging in the vectors for x_1 and x_2 into the formulas for s, w and b we get

$$s = \frac{-1}{31} \approx -0.3226$$
$$w \approx (-0.1935, -0.0323, 0.1613)$$
$$b \approx -0.2581.$$

The reason why we add s is that otherwise, the system of equations would not be solvable.

4 5 dims I

The procedure for the previous problem works here as well, only with the definitions of x_1 and x_2 now changed. Plugging these vectors into the formula for s, w and b, we get

$$s \approx 0.0211$$
$$w \approx (-0.0421, -0.0211, 0.0842, -0.1474, 0.1053)$$
$$b \approx 0.03158.$$

5 5 dims II

II **Answer for yourself:**

- what is a possible mathematical criterion to test that w is not parallel to the line $x_2 - x_3$?

  Solution: Check the vector $(x_2 - x_3)/w$, where / is here meant element-wise. The resulting vector is constant if and only if w is parallel to $x_2 - x_3$.

- what is a possible mathematical criterion to test that w is not orthogonal to the line $x_1 - x_3$?

  Solution: Check $w\cdot (x_1 - x_3) \neq 0$.

III

Solution: Let us choose an initial vector $w = x_2 - x_3 + e_2 = (1, 3, -3, 1, -1)$, by adding e_2 we are ensuring that w is not parallel to $x_2 - x_3$.

Now we can orthogonalise w:

$$\begin{aligned} w_2 &= w - (w \cdot z)\frac{z}{\|z\|^2} \\ &= \frac{1}{8}(-1, 6, 3, -1, 1) \end{aligned}$$

where $z = x_2 - x_3 = (1, 2, -3, 1, -1)$.

IV

Solution: We can now momentarily forget about x_3 since $f(x_2) = f(x_3)$ with our choice of weights w_2. Now we want $f(x_1) = 1$ and $f(x_2) = -1$ for some scaled weight $w_3 = sw_2$ and bias b.

The resulting linear system can be solved by

$$\begin{aligned} s &= \frac{2}{w_2 \cdot (x_1 - x_2)} \approx 1.7777 \\ b &= -\frac{1}{2} s w_2 (x_1 + x_2) \approx 1.222 \end{aligned}$$

so that

$$w_3 \approx (-0.2222, 1.3333, 0.6666, -0.2222, 0.2222).$$

V

Solution
Consider the figure below, where $x_1 - x_2$ is orthogonal to $x_2 - x_3 = w$. There is no way to define a line that separates x_1 from x_2, x_3 and is orthogonal to $x_2 - x_3$, since any line orthogonal to $x_2 - x_3$ will have x_2 and x_1 on the same side.