Lucía Rivas

ADVANCED COURSE IN MACHINE LEARNING : 2

(4) Automatic differentiation : $f(w,x) = \dfrac{w_1x_1 + w_2x_2}{(w_1x_1 + w_2x_2)^3} + w_1^2$

Variable	Computation	Derivative
w_3	$w_1 x_1$	$dw_3/dw_1 = x_1$
w_4	$w_2 x_2$	$dw_4/dw_2 = x_2$
w_5	$w_3 + w_4$	$dw_5/dw_i = 1$ for $i = 1, 2$
w_6	w_5^3	$dw_6/dw_5 = 3w_5^2$
w_7	$\dfrac{w_5}{w_6}$	$dw_7/dw_5 = \dfrac{1}{w_6}$, $dw_7/dw_6 = \dfrac{-w_5}{w_6^2}$
w_8	w_1^2	$dw_8/dw_1 = 2w_1$
w_9	$w_7 + w_8$	$dw_9/dw_7 = 1$, $dw_9/dw_8 = 1$

Now we go backwards for $f(w,x) = w_9$

- $\dfrac{df}{dw_9} = 1$
- $\dfrac{df}{dw_8} = \dfrac{df}{dw_9} \cdot \dfrac{dw_9}{dw_8} = 1$
- $\dfrac{df}{dw_7} = \dfrac{df}{dw_9} \cdot \dfrac{dw_9}{dw_7} = 1$
- $\dfrac{df}{dw_6} = \dfrac{df}{dw_7} \cdot \dfrac{dw_7}{dw_6} = -\dfrac{w_5}{w_6^2}$
- $\dfrac{df}{dw_5} = \dfrac{df}{dw_7} \cdot \dfrac{dw_7}{dw_5} + \dfrac{df}{dw_6} \cdot \dfrac{dw_6}{dw_5} = \dfrac{1}{w_6} - \dfrac{w_5}{w_6^2} 3w_5^2 = \dfrac{w_6 - 3w_5^3}{w_6^2}$
- $\dfrac{df}{dw_4} = \dfrac{df}{dw_5} \cdot \dfrac{dw_5}{dw_4} = \dfrac{w_6 - 3w_5^3}{w_6^2}$
- $\dfrac{df}{dw_3} = \dfrac{df}{dw_5} \cdot \dfrac{dw_5}{dw_3} = \dfrac{w_6 - 3w_5^3}{w_6^2}$

SOLUTION

- $\boxed{\dfrac{df}{dw_2}} = \dfrac{df}{dw_4} \cdot \dfrac{dw_4}{dw_2} = x_2\left(\dfrac{w_6 - 3w_5^3}{w_6^2}\right) = x_2 \dfrac{(w_1x_1 + x_2w_2)^3 - 3(w_1x_1 + w_2x_2)^3}{(w_1x_1 + w_2x_2)^6}$
- $\boxed{\dfrac{df}{dw_1}} = \dfrac{df}{dw_3} \cdot \dfrac{dw_3}{dw_1} + \dfrac{df}{dw_8} \cdot \dfrac{dw_8}{dw_1} = x_1 \dfrac{w_6 - 3w_5^3}{w_6^2} + 2w_1 =$

$$= x_1 \dfrac{(w_1x_1 + w_2x_2)^3 - 3(w_1x_1 + w_2x_2)^3}{(w_1x_1 + w_2x_2)^6} + 2w_1$$

① One-dim regression with model $\bar{y} = \alpha x$, loss $L(y, x, \alpha) = (y - \bar{y})^2$

a) Write the def. of risk $R(\alpha)$ and compute it. α for smallest R?

The risk is defined as the expected loss over the distribution.

$$R(\alpha) = E_{p(x,y)}[L(y,x,\alpha)] = \int_{x,y} L(y,x,\alpha)\, p(x,y)\, dy\, dx$$

We know $L(y,x,\alpha)$, $p(x)$ and $p(y|x)$ so we can use the definition of conditional prob: $p(x,y) = p(y|x) \cdot p(x)$

$$p(x) = \begin{cases} 1/6, & x \in [-3,3] \\ 0 & \text{otherwise} \end{cases} \qquad p(y|x) = \begin{cases} 1, & y \in [2x - 0'5,\ 2x + 0'5] \\ 0 & \text{otherwise} \end{cases}$$

$$R(\alpha) = \int_{x,y} L(y,x,\alpha)\, p(x)\, p(y|x)\, dy\, dx = \int_{-3}^{3} \int_{2x-0'5}^{2x+0'5} (y - \alpha x)^2 \cdot \frac{1}{6}\, dy\, dx =$$

$$= \int_{-3}^{3} \frac{(y-\alpha x)^3}{18} \Big|_{2x-0'5}^{2x+0'5} dx = \frac{1}{18} \int_{-3}^{3} \left(((2-\alpha)x + 0'5)^3 - ((2-\alpha)x - 0'5) \right)^3 dx =$$

$$= \frac{1}{18} \int_{-3}^{3} \left(3(2-\alpha)^2 x^2 + \frac{1}{4} \right) dx = \frac{(2-\alpha)^2 x^3}{18} + \frac{x}{72} \Big|_{-3}^{3} =$$

$$= \frac{(2-\alpha)^2\, 54}{18} + \frac{6}{72}$$

It is easy to see that $R(\alpha)$ is min when $\alpha = 2$.

③ Lagrange multipliers for $\min\ x^2 + y^2$ st. $3x - y \leq -2$

a) Lagragian and requirements for λ

Lagragian: $L(\theta, \lambda) = f(\theta) + \lambda g(\theta)$ with $\theta = (x, y)$

and $f(x,y) = x^2 + y^2$, $g(x,y) = 3x - y + 2$ so

$L(x, y, \lambda) = x^2 + y^2 + \lambda(3x - y + 2)$

Condition for λ: $\lambda \geq 0$.

b) Find the optimal solution.

$$\nabla L(x,y,\lambda) = (2x + 3\lambda,\ 2y - \lambda,\ -y + 3x + 2) = \bar{0}$$

$$\begin{cases} 2x + 3\lambda = 0 \\ 2y - \lambda = 0 \\ -y + 3x + 2 = 0 \end{cases}$$ Solving it we have: $\boxed{(x, y, \lambda) = (-3/5,\ 2/5,\ 1/5)}$

c) Iterative solution

1°) Let $\lambda = 0$ $\Rightarrow$ $L(x, y, 0) = x^2 + y^2$

$\nabla L(x, y) = (2x, 2y, 0) = (0, 0, 0)$

so we have $(x, y) = (0, 0)$ but it does'nt check the inequality $3x - y \leq -2$, $0 \nleq -2$

we move λ up.

2°) $\lambda = 1 \Rightarrow \nabla L(x, y) = (2x + 3, 2y - 1)$

$(x, y) = (-3/2, 1/2)$

$-\frac{9}{2} - \frac{1}{2} \leq -2$ ✓ so $\lambda = \lambda - \frac{1}{2}$

3°) $\lambda = \frac{1}{2} \Rightarrow \nabla L(x, y) = (2x + 3/2, 2y - \frac{1}{2})$

$(x, y) = (-3/4, 1/4)$

$-\frac{9}{4} - \frac{1}{4} \leq -2$ ✓ so $\lambda = \lambda - \frac{1}{4}$

4°) $\lambda = \frac{1}{4} \Rightarrow \nabla L(x, y) = (2x + \frac{3}{4}, 2y - \frac{1}{4})$

$(x, y) = (-3/8, 1/8)$

$-9/8 - 1/8 \leq -2$ NO so $\lambda = \lambda + \frac{1}{8}$

5°) $\lambda = 3/8 \Rightarrow \nabla L(x, y) = (2x + \frac{9}{8}, 2y - \frac{3}{8})$

$(x, y) = (-9/16, 3/16)$

$\frac{-27 - 3}{16} \leq -2$ NO so $\lambda = \lambda + \frac{1}{16}$

and so on...