ICTS 4156: Introduction to ML

Razvan Bunescu

Lecture notes, February 3, 2021

1 Notes on lecture slides material

Let $f(x,y) = (2x+3y-1)^2 + (x-2y+1)^2$. The multivariate function $f(x,y)$ is minimized when the gradient is equal with 0. Setting the gradient (derivative) to 0 results in:

$$
\begin{aligned}
\frac{\delta f}{\delta x} = 0 \quad &=> \quad 2*(2x+3y-1)*2+2*(x-2y+1)*1 = 0 \mid \text{divide by } 2 && (1)\\
&=> \quad 4x+6y-2+x-2y+1 = 0 && (2)\\
&=> \quad 5x+4y-1 = 0 && (3)\\
\frac{\delta f}{\delta y} = 0 \quad &=> \quad 2(2x+3y-1)*3+2*(x-2y+1)*-2 = 0 \mid \text{divide by } 2 && (4)\\
&=> \quad 6x+9y-3-2x+4y-2 = 0 && (5)\\
&=> \quad 4x+13y-5 = 0 && (6)
\end{aligned}
$$

Solve it using variable substitution. Can write code for it, or use linear algebra functions from `NumPy`, i.e. `numpy.linalg.solve`.

In the homework assignment, you will similarly need to compute the gradient of $J(\mathbf{w})$ and set it to 0, where:

$$
\begin{aligned}
J(\mathbf{w}) &= \frac{1}{2N}\sum_{n=1}^{N}(w_1 x_n + w_0 - t_n)^2 && (7)\\
\frac{\delta J}{\delta w_1} &= 0 => \ldots && (8)\\
\frac{\delta J}{\delta w_0} &= 0 => \ldots && (9)
\end{aligned}
$$

2 Quiz week 2

2.1 Bayes rule

Problem: There are 2 bowls of fruits: bowl B1 contains 1 apple and 1 orange, while bowl B2 contains 2 apples and 3 oranges. Patrick is blindfolded, so he first picks one bowl uniformly at random (i.e. 0.5 probability to pick each of them), then picks one fruit from the bowl uniformly at random (i.e. all fruits are equally likely to be chosen). If Patrick picked an apple, what is the probability that it came from bowl B1?

Solution: We need to computer $P(B_1|a)$. By Bayes Rule, we have:

$$P(B_1|a) = \frac{P(a|B_1)P(B_1)}{P(a)}$$

By the Sum Rule of probability and by the definition of conditional probabilities, we have that $P(a)$ can be computed as:

$$\begin{aligned} P(a) &= P(a, B_1) + P(a, B_2) \\ &= P(a|B_1)P(B_1) + P(a|B_2)P(B_2) \\ &= 1/2 * 1/2 + 2/5 * 1/2 \\ &= 1/2(1/2 + 2/5) \\ &= 1/2 * 9/10 \end{aligned}$$

Plugging this in the Bayes Rule formula above we get:

$$\begin{aligned} P(B_1|a) &= \frac{1/2 * 1/2}{1/2 * 9/10} \\ &= 1/2 * 10/9 \\ &= 5/9 \end{aligned}$$

3 Differentiation rules and examples

$$
\begin{aligned}
f(x) = 2x \quad &=> \quad \frac{\delta f}{\delta x} = 2 \\
f(x) = x^2 \quad &=> \quad \frac{\delta f}{\delta x} = 2x \\
f(x) = x^k \quad &=> \quad \frac{\delta f}{\delta x} = kx^{k-1} \\
h(x) = f(g(x)) \quad &=> \quad \frac{\delta h}{\delta x} = \frac{\delta f}{\delta g}\frac{\delta g}{\delta x} \\
h(x) = (g(x))^2 \quad &=> \quad \frac{\delta h}{\delta x} = 2g(x)\frac{\delta g}{\delta x} \\
f(g) = g^2 \quad &=> \quad \frac{\delta f}{\delta g} = 2g \\
f(x) = g(x) + h(x) \quad &=> \quad \frac{\delta f}{\delta x} = \frac{\delta g}{\delta x} + \frac{\delta h}{\delta x} \\
f(x) = g(x)h(x) \quad &=> \quad \frac{\delta f}{\delta x} = \frac{\delta g}{\delta x}h(x) + g(x)\frac{\delta h}{\delta x} \\
f(x) = \frac{g(x)}{h(x)} \quad &=> \quad \frac{\delta f}{\delta x} = \frac{g'(x)h(x) - h'(x)g(x)}{h^2(x)} \\
f(x) = \ln(x) \quad &=> \quad \frac{\delta f}{\delta x} = 1/x \\
f(x) = \ln(2x+1) \quad &=> \quad \frac{\delta f}{\delta x} = \frac{2}{2x+1} \\
f(x) = e^x \quad &=> \quad \frac{\delta f}{\delta x} = e^x \\
f(x) = e^{2x+1} \quad &=> \quad \frac{\delta f}{\delta x} = 2e^{2x+1} \\
f(x,y) = (2x+3y+1)^2 \quad &=> \quad \frac{\delta f}{\delta x} = 2(2x+3y+1)2 \\
&\quad\quad\;\; \frac{\delta f}{\delta y} = 2(2x+3y+1)3 \\
&\quad\quad\;\; \nabla f = [\frac{\delta f}{\delta x}, \frac{\delta f}{\delta y}] = ...
\end{aligned}
$$