AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCES

(AIMS RWANDA, KIGALI)

Name: Darix SAMANI SIEWE Assignment Number: 1

Course: Data-Driven Optimization Date: February 2, 2025

1. Show the following.

   (a) for sigmoid function $\sigma t = \frac{1}{1+e^{-t}}$ show that its derivative is $\sigma^{'}(t) = \sigma^{'}(t)[1 - \sigma(t)]$

$$
\begin{aligned}
\sigma^{'}(t) &= \frac{e^{-t}}{(1+e^{-t})^2} = \frac{e^{-t}+1-1}{(1+e^{-t})^2} \\
&= \frac{1}{1+e^{-t}}[1 - \frac{1}{1+e^{-t}}] \\
&= \sigma(t)[1 - \sigma(t)]
\end{aligned}
$$

   (b) If g : $R^n \rightarrow$ R and $g(x) = log(\sigma(f(x)))$, then show that $\nabla g(x) = [1 - \sigma(f(x))]\nabla f(x)$

$$
\begin{aligned}
\nabla g(x) &= \frac{\nabla \sigma f(x)}{\sigma f(x)} \\
&= \frac{\sigma^{'}(f(x))\nabla f(x)}{\sigma f(x)} \\
&= \sigma(f(x))[1 - \sigma(f(x))] \\
&= [1 - \sigma(f(x))]\nabla f(x)
\end{aligned}
$$

   (c) (left as an exercise to the curious) Show that the Hessian matrix of $g(x) = log(\sigma(f(x)))$ (defined in 3b)) is given by show that the hessian matrix of $g(x) = log(\sigma f(x))$

$$
\begin{aligned}
\nabla^2 g(x) &= \nabla(\nabla g(x)) = \nabla\left[(1 - \sigma(f(x)))\nabla f(x)\right] \\
&= \nabla\left[1 - \sigma(f(x))\right]\nabla f(x) + \left[1 - \sigma(f(x))\right]\nabla^2 f(x) \\
&= -\sigma'(f(x))\nabla f(x)\nabla f(x)^T + (1 - \sigma(f(x)))\nabla^2 f(x) \\
&- -[\sigma(f(x))][1 - \nabla f(x)]\nabla f(x)\nabla f(x)^T + (1 - \sigma(f(x)))\nabla^2 f(x) \\
&= [1 - \nabla f(x)][\nabla^2 f(x) - \sigma(f(x))]\nabla f(x)\nabla f(x)^T
\end{aligned}
$$