Sigmoid Operator

Rylan W. Yancey

10/1/2023

Definition & Gradient

The Sigmoid function is defined as $\sigma(x) = \frac{1}{1+e^{-x}}$ To find the gradient, we will apply the chain rule and simplify.

$$
\begin{aligned}
\frac{d}{dx}\sigma(x) &= \frac{d}{dx}\left[\frac{1}{1+e^{-x}}\right] && (1)\\
&= \frac{d}{dx}\left(1+\mathrm{e}^{-x}\right)^{-1} && (2)\\
&= -(1+e^{-x})^{-2}(-e^{-x}) && (3)\\
&= \frac{e^{-x}}{\left(1+e^{-x}\right)^{2}} && (4)\\
&= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} && (5)\\
&= \frac{1}{1+e^{-x}} \cdot \frac{(1+e^{-x})-1}{1+e^{-x}} && (6)\\
&= \frac{1}{1+e^{-x}} \cdot \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}}\right) && (7)\\
&= \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right) && (8)\\
&= \sigma(x) \cdot (1-\sigma(x)) && (9)
\end{aligned}
$$

Therefore, we say that the gradient of the sigmoid function with respect to x is $\sigma(x) \cdot (1-\sigma(x))$.

Contributions

Thanks to Michael Percy for the detailed steps.