OTH Amberg-Weiden
Fakultät für Elektrotechnik, Medien und Informatik
Prof. Dr.-Ing. Christian Bergler, Prof. Dr. Fabian Brunner

Ostbayerische
Technische Hochschule
Amberg-Weiden

Deep Learning – Summer Semester 2024
Exercise for Lecture – Mathematical Foundation

Task 1 (Differentiation)
Calculate the first-order derivative $f'(x)$ in each case:

a) $f(x) = 5x^4$

b) $g(x) = 3\sin(x)^2$

c) $z(x) = \log(\sqrt{e^{2x}})$

Task 2 (Tangent Hyperbolicus)
The hyperbolic tangent

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

is an example of a (non-linear) activation function frequently used in deep learning applications.

a) Draw the graph of the function

b) Calculate the first-order derivative $\tanh'(x)$

c) Show that $\tanh'(x) = 1 - (\tanh(x))^2$

Task 3 (Sigmoid Activation Function)
The sigmoid function

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

is an example of a (non-linear) activation function frequently used in deep learning applications.

a) Draw the graph of the function

b) Calculate the first-order derivative $\sigma'(x)$

c) Show that $\sigma'(x) = \sigma(x)(1 - \sigma(x))$

Task 4 (Convex Functions)

a) Decide whether the following functions are convex or not:

$$f(x) = |x|\,, \quad g(x) = \sin(x)\,, \quad h(x) = e^x$$

b) Let f and g be convex functions. Show that $h(x) := \max(f(x), g(x))$ is also convex

c) Conclude from subtask a) that the ReLU activation function $\mathrm{ReLU}(x) = \max(0, x)$ is convex

d) Let f and g be convex functions. Show that $f + g$ is also convex

e) Let f be a convex function and h a linear function. Show that $f \circ h$ is convex.

Task 5 (Logistic Regression)
The loss function in logistic regression has the form

$$L(\boldsymbol{\theta}) = -\sum_{i=1}^{m} y^{(i)} \log(\sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)})) + (1 - y^{(i)}) \log(1 - \sigma(\boldsymbol{\theta}^T \boldsymbol{x}^{(i)}))$$

where $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function and $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_p)^T$ denotes the weights of the model. The convention $\boldsymbol{x}_0^{(i)} := 1$ is used.

a) Show that $f_1(x) = -\log(\sigma(x))$ and $f_2(x) = -\log(1 - \sigma(x))$ are convex. Calculate their second derivatives

b) Conclude from a) that L is a convex function in the weights $\boldsymbol{\theta}$

c) Calculate the partial derivatives of L according to the weights $\theta_0, \ldots, \theta_p$ and describe the gradient $\nabla L_{\boldsymbol{\theta}}$ in vectorized form. Use the representation from task 3c

Task 1

a) $f(x) = 5x^4$

$f'(x) = 20x^3$

b) $g(x) = 3\sin(x)^2$

$g'(x) = 6\cos(x)$

c) $z(x) = \log(\sqrt{e^{2x}})$

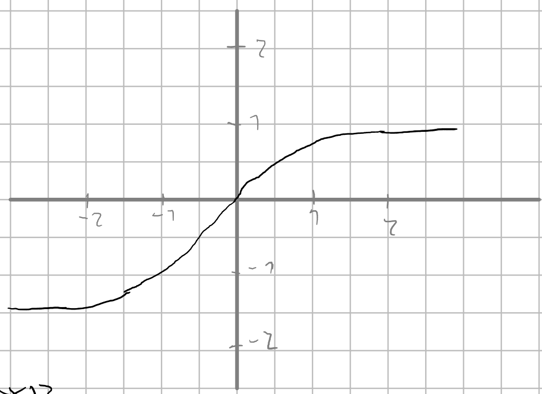
$z'(x) = \dfrac{\log(\sqrt{e^{2x}})}{\log(10)}$

$z'(x) = \dfrac{\frac{1}{\cancel{\sqrt{e^{2x}}}} \cdot \cancel{\frac{1}{2}}\cancel{(e^{2x})^{-\frac{1}{2}}} \cdot \cancel{2}}{\log(10)}$

$z'(x) = \dfrac{1}{\log(10)}$

Task 2

$\tanh(x) = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$

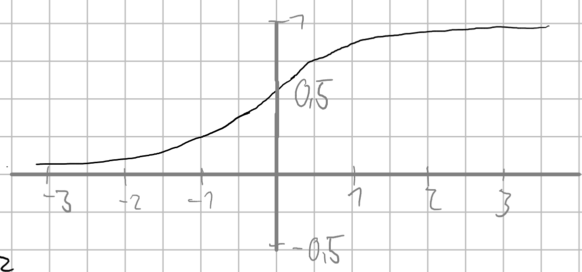


$\tanh'(x) = \dfrac{(e^x + e^{-x})\cdot(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2}$

$= \dfrac{(e^x + e^{-x})^2 - (e^x - e^{-x})^2}{(e^x + e^{-x})^2} = 1 - \dfrac{(e^x - e^{-x})^2}{(e^x + e^{-x})^2} = 1 - (\tanh(x))^2$

c)*

Task 3



$\sigma(x) = \dfrac{1}{1+e^{-x}}$ $\left(\left(\frac{1}{u(x)}\right)' = -\frac{u'(x)}{u(x)^2}\right)$

$\sigma'(x) = -\dfrac{(1+e^{-x})'}{(1+e^{-x})^2} = -\dfrac{0 + (-e^{-x})}{(1+e^{-x})^2} = -\dfrac{-e^{-x}}{(1+e^{-x})^2}$

$= \dfrac{e^{-x}}{(1+e^{-x})^2} = \dfrac{1}{1+e^{-x}} \cdot \left(1 - \dfrac{1}{1+e^{-x}}\right) = \sigma(x)\cdot(1-\sigma(x))$

c)*2

Task 4

$g(x) = \sin(x)$

let $x_1 = \frac{\pi}{2}$, $x_2 = \frac{3\pi}{2}$, $x_3 = \pi$

since $g(x_1) = -1$, $g(x_2) = -1$ and $g(x_3) = 0$

$g(x)$ cannot be convex

$h(x) = e^x$

since $h''(x) = e^x > 0 \ \forall x$

$h(x)$ is convex

*Task 2c

$$\tanh(x) = \int \tanh'(x)$$

$$\tanh(x) = \int 1 - (\tanh(x))^2 \, dx \qquad \tanh(x)^2 + \operatorname{sech}(x)^2 = 1$$

$$= \int \operatorname{sech}(x)^2 \, dx$$

$$= \tanh(x) + C$$

*² Task 3c

$$\sigma(x) = \int \sigma(x)(1 - \sigma(x)) \, dx$$

$$= \int \sigma(x) - \sigma(x)^2 \, dx$$

$$= \int \frac{1}{1+e^{-x}} - \left(\frac{1}{1+e^{-x}}\right)^2 dx$$

$$= \int \frac{1 - \frac{1}{1+e^{-x}}}{(1+e^{-x})} \, dx$$

$$= \int \frac{e^x}{(e^x+1)^2} \, dx \qquad | \, u = e^x + 1 \rightarrow du = e^x dx$$

$$= \int \frac{1}{u^2} \, du$$

$$= -\frac{1}{u}$$

$$= -\frac{1}{e^x+1} = \frac{1}{e^{-x}+1}$$

Task 5

a) Show $f_1 = -\log(\sigma(x))$ and $f_2(x) = -\log(1 - \sigma(x))$ are convex

with $\sigma(x) = \frac{1}{1+e^{-x}}$

strict convex

$$f_1(x) = -\log\left(\frac{1}{1+e^{-x}}\right) = -(\underbrace{\cancel{\log}(1)}_{=0} - \log(1+e^{-x})) = \log(1+e^{-x})$$

log'

$$f_1'(x) = -\left(\frac{1}{\frac{1}{1+e^{-x}}}\right) \cdot -(1+e^{-x})^{-2} \cdot (-e^{-x}) = -\frac{e^{-x}}{1+e^{-x}}$$

nach diff.

$$f_1''(x) = \left(\frac{1}{1+e^{-x}} \cdot (-e^{-x})\right) = -\underbrace{(1+e^{-x})^{-2} \cdot (-e^{-x})}_{\text{nachdiff.}} \cdot (-e^{-x}) + \frac{1}{1+e^{-x}} \cdot e^{-x}$$

„mal 1"

$$= -\frac{(-e^{-x}) \cdot (-e^{-x})}{(1+e^{-x})^2} + \frac{e^{-x}}{1+e^{-x}} \cdot \frac{1+e^{-x}}{1+e^{-x}}$$

$$= -\frac{(-e^{-x})\cdot(-e^{-x})}{(1+e^{-x})^2} + \frac{e^{-x} + e^{-2x}}{(1+e^{-x})^2}$$

$$= -\frac{e^{-2x} + e^{-x} + e^{-2x}}{(1+e^{-x})^2} = \frac{e^{-x}}{(1+e^{-x})^2}$$

$$f_2(x) = -\log(1 - \sigma(x)) = -\log\left(1 - \frac{1}{1+e^{-x}}\right)$$

rest siehe Lösung