



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 $\text{ReLU}(x) = \max(0, x)$ 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

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

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

The convention $x_0^{(i)} := 1$ 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 $\theta_0, \dots, \theta_p$ and describe the gradient $\nabla L_{\boldsymbol{\theta}}$ in vectorized form. Use the representation from task 3c

Task 1

$$\begin{aligned} \text{a) } f(x) &= 5x^4 \\ f'(x) &= 20x^3 \end{aligned}$$

$$\begin{aligned} \text{b) } g(x) &= 3 \sin(x)^2 \\ g'(x) &= 6 \cos(x) \end{aligned}$$

$$\begin{aligned} \text{c) } z(x) &= \log(\sqrt{e^{2x}}) \\ z'(x) &= \frac{\log(\sqrt{e^{2x}})}{\log(10)} \\ z'(x) &= \frac{\frac{1}{\sqrt{e^{2x}}} \cdot \frac{1}{2} (e^{2x})^{-\frac{1}{2}} \cdot 2}{\log(10)} \\ z'(x) &= \frac{1}{\log(10)} \end{aligned}$$

Task 2

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

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

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