

**Begonnen am** Monday, 1. February 2021, 14:07

**Status** Beendet

**Beendet am** Monday, 1. February 2021, 14:20

**Verbrauchte Zeit** 13 Minuten 10 Sekunden

**Bewertung** 8,00 von 10,00 (80%)

**Frage 1**

Falsch

Erreichte Punkte  
0,00 von 1,00

How would you implement the inner product between two (1D) numpy arrays **a** and **b**? You can assume `import numpy as np`.

Antwort:



Die richtige Antwort ist: `a @ b`

**Frage 2**

Richtig

Erreichte Punkte  
1,00 von 1,00

What is the likelihood (given a single sample),  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y})$ , of a typical logistic regression model,  $g(\mathbf{x}; \theta)$ ?

Wählen Sie eine Antwort:

- ☐ a.  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = g(\mathbf{x}; \theta) \cdot \mathbf{y} + (g(\mathbf{x}; \theta) - 1) \cdot (\mathbf{y} - 1)$
- ☐ b.  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = g(\mathbf{x}; \theta)^{\mathbf{y}} + (g(\mathbf{x}; \theta) - 1)^{\mathbf{y}-1}$
- ☒ c.  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = g(\mathbf{x}; \theta)^{\mathbf{y}} + (1 - g(\mathbf{x}; \theta))^{1-\mathbf{y}}$   
✓
- ☐ d.  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = \mathbf{y}^{g(\mathbf{x}; \theta)} + (1 - \mathbf{y})^{1-g(\mathbf{x}; \theta)}$
- ☐ e.  $\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = \mathbf{y}^{g(\mathbf{x}; \theta)} + (\mathbf{y} - 1)^{g(\mathbf{x}; \theta)-1}$

Your answer is correct.

Die richtige Antwort lautet:

$$\mathcal{L}(\theta; \mathbf{x}, \mathbf{y}) = g(\mathbf{x}; \theta)^{\mathbf{y}} + (1 - g(\mathbf{x}; \theta))^{1-\mathbf{y}}$$

**Frage 3**

Richtig

Erreichte Punkte  
1,00 von 1,00

What is the one-hot encoding of the label 3, given a set of integer labels 0, 1, 2, 3, 4, 5 (assuming natural ordering)? The answer is assumed to be 6 numbers without whitespace or commas in between.

Antwort:



Die richtige Antwort ist: 000100

## Frage 4

Richtig

Erreichte Punkte  
1,00 von 1,00

What is the negative log-likelihood (given a single sample),  $-l(\theta; \mathbf{x}, \mathbf{y})$ , for a typical linear regression model,  $g(\mathbf{x}; \theta)$ , where the noise is assumed to have variance  $\sigma^2$ ?

Wählen Sie eine Antwort:

- ☐ a.  $-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma) + \frac{g(\mathbf{x}\theta) - \mathbf{y}}{2\sigma}$
- ☒ b.  $-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma^2) + \frac{(g(\mathbf{x}\theta) - \mathbf{y})^2}{2\sigma^2}$  ✓
- ☐ c.  $-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma) + \frac{\mathbf{y} - g(\mathbf{x}\theta)}{2\sigma}$
- ☐ d.  $-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma) + \frac{g(\mathbf{x}\theta) + \mathbf{y}}{2\sigma}$
- ☐ e.  $-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma^2) + \frac{(g(\mathbf{x}\theta) + \mathbf{y})^2}{2\sigma^2}$

Your answer is correct.

Die richtige Antwort lautet:

$$-l(\theta; \mathbf{x}, \mathbf{y}) = \frac{1}{2} \ln(2\pi\sigma^2) + \frac{(g(\mathbf{x}\theta) - \mathbf{y})^2}{2\sigma^2}$$

## Frage 5

Richtig

Erreichte Punkte  
1,00 von 1,00

What formula did we use to implement the gradient check for `numpy` gradients? Hint: it is the same formula as the one you used for the simple gradient checking exercise.

Wählen Sie eine Antwort:

- ☐ a.  $\frac{f(x+h) - f(x)}{h}$
- ☐ b.  $\frac{f(x+h) - f(x-h)}{h}$
- ☒ c.  $\frac{f(x+h) - f(x-h)}{2h}$  ✓
- ☐ d.  $\frac{f(x) + f(x-h)}{h}$
- ☐ e.  $\frac{f(x+h) + f(x-h)}{2h}$

Your answer is correct.

Die richtige Antwort lautet:  $\frac{f(x+h) - f(x-h)}{2h}$ 

## Frage 6

Falsch

Erreichte Punkte  
0,00 von 1,00

What do we use numerical differentiation for in `numpy`?

Wählen Sie eine oder mehrere Antworten:

- ☐ a. Testing the gradients during development.
- ☒ b. Computing the gradients for updating the network. ✗
- ☐ c. Checking the statistics of the gradients computed.
- ☐ d. Finding errors in the gradient computation of a fully trained network.
- ☐ e. `numpy` does not make use of numerical differentiation.

Your answer is incorrect.

Die richtige Antwort lautet: Testing the gradients during development.

**Frage 7**

Richtig

Erreichte Punkte  
1,00 von 1,00

How would you describe the computations of the gradients of a convolutional layer w.r.t. inputs, if the forward pass is implemented by means of convolution (**not** cross-correlation).

Wählen Sie eine Antwort:

- ☒ a. The cross-correlation between gradients and kernel weights ✓
- ☐ b. The convolution between gradients and inputs
- ☐ c. The cross-correlation between gradients and inputs
- ☐ d. The convolution between gradients and kernel weights
- ☐ e. The sum of gradients

Your answer is correct.

Die richtige Antwort lautet: The cross-correlation between gradients and kernel weights

**Frage 8**

Richtig

Erreichte Punkte  
2,00 von 2,00

How would you describe the computations of the gradients of a convolutional layer w.r.t. kernel weights, if the forward pass is implemented by means of convolution (**not** cross-correlation).

Wählen Sie eine Antwort:

- ☐ a. The convolution between gradients and kernel weights
- ☒ b. The convolution between inputs and gradients ✓
- ☐ c. The cross-correlation between gradients and kernel weights
- ☐ d. The cross-correlation between inputs and gradients
- ☐ e. The sum of gradients

Your answer is correct.

Die richtige Antwort lautet: The convolution between inputs and gradients

**Frage 9**

Richtig

Erreichte Punkte  
1,00 von 1,00

What is considered to be the state of an Adam optimiser **with bias correction**?

Wählen Sie eine Antwort:

- ☒ a. the number of steps and the moving average of gradients and squared gradients ✓
- ☐ b. the moving average of gradients and squared gradients
- ☐ c. the moving average of squared gradients
- ☐ d. the moving average of gradients
- ☐ e. the moving average of parameters

Your answer is correct.

Die richtige Antwort lautet: the number of steps and the moving average of gradients and squared gradients

