## Exercise 1 - Eigenvalues and Eigenvectors

You are given the following set of eigevalues and eigenvectors. Compute the corresponding matrix.

 $\lambda_1 = 1, \, \lambda_2 = 2, \, \mathbf{v}_1 = (\sqrt{0.5}, \sqrt{0.5})^\top, \, \mathbf{v}_2 = (\sqrt{0.5}, -\sqrt{0.5})^\top.$ 

## **Exercise 2 - Parameter Counting**

Use PyTorch to load the **alexnet** model and automatically compute its number of parameters. Output the number of parameters for each layer and the total number of parameters in the model.

## **Exercise 3 - Convolutional Layers**

Consider the following  $4 \times 4 \times 1$  input X and a  $2 \times 2 \times 1$  convolutional kernel K with no bias term

$$X = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 1 & -1 & 0 & 1 \end{pmatrix}, \qquad K = \begin{pmatrix} 1, & 2 \\ 0, & 1 \end{pmatrix}$$

- (a) What is the output of the convolutional layer for the case of stride 1 and no padding?
- (b) What if we have stride 2 and no padding?
- (c) What if we have stride 2 and zero-padding of size 1?

## **Exercise 4 - Scaled Dot-Product Attention**

Consider the matrices Q, K, V given by

$$Q = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & 2 \\ 0 & 3 & -1 \end{pmatrix}.$$

Compute the context matrix C using the scaled dot product attention.