Deep Learning

Theoretical Exercises – Week 1 – Chapter 2

Exercises on the book ”Deep Learning” written by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.
Exercises and solutions by T. Méndez and G. Schuster

FS 2024

1 Exercises on Linear Algebra

1. Calculate the following matrix products.

   (a) $\boldsymbol{C} = \boldsymbol{AB}$ and $\boldsymbol{D} = \boldsymbol{BA}$

   with $\boldsymbol{A} = \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix}$ and $\boldsymbol{B} = \begin{pmatrix} 7 & 0 \\ 0 & 1 \\ 3 & 8 \end{pmatrix}$

   Solution:
   Simple matrix multiplication (row times column):

$$\boldsymbol{C} = \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix} \begin{pmatrix} 7 & 0 \\ 0 & 1 \\ 3 & 8 \end{pmatrix} = \begin{pmatrix} 7 \cdot 7 + 7 \cdot 0 + 0 \cdot 3 & 7 \cdot 0 + 7 \cdot 1 + 0 \cdot 8 \\ 8 \cdot 7 + 4 \cdot 0 + 2 \cdot 3 & 8 \cdot 0 + 4 \cdot 1 + 2 \cdot 8 \end{pmatrix} = \begin{pmatrix} 49 & 7 \\ 62 & 20 \end{pmatrix}$$

$$\boldsymbol{D} = \begin{pmatrix} 7 & 0 \\ 0 & 1 \\ 3 & 8 \end{pmatrix} \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 7 \cdot 7 + 0 \cdot 8 & 7 \cdot 7 + 0 \cdot 4 & 7 \cdot 0 + 0 \cdot 2 \\ 0 \cdot 7 + 1 \cdot 8 & 0 \cdot 7 + 1 \cdot 4 & 0 \cdot 0 + 1 \cdot 2 \\ 3 \cdot 7 + 8 \cdot 8 & 3 \cdot 7 + 8 \cdot 4 & 3 \cdot 0 + 8 \cdot 2 \end{pmatrix} = \begin{pmatrix} 49 & 49 & 0 \\ 8 & 4 & 2 \\ 85 & 53 & 16 \end{pmatrix}$$

   (b) $\boldsymbol{b} = \boldsymbol{Ax}$

   with $\boldsymbol{A} = \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix}$ and $\boldsymbol{x} = \begin{pmatrix} 7 \\ 3 \\ 10 \end{pmatrix}$

   Solution:
   Simple matrix multiplication (row times column):

$$\boldsymbol{b} = \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix} \begin{pmatrix} 7 \\ 3 \\ 10 \end{pmatrix} = \begin{pmatrix} 7 \cdot 7 + 7 \cdot 3 + 0 \cdot 10 \\ 8 \cdot 7 + 4 \cdot 3 + 2 \cdot 10 \end{pmatrix} = \begin{pmatrix} 70 \\ 88 \end{pmatrix}$$

(c) $\boldsymbol{c} = \boldsymbol{h}^T\boldsymbol{A}$

with $\boldsymbol{A} = \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix}$ and $\boldsymbol{h} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$

Solution:

Simple matrix multiplication (row times column):

$$\boldsymbol{c} = \begin{pmatrix} 2 & 5 \end{pmatrix} \begin{pmatrix} 7 & 7 & 0 \\ 8 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 7 + 5 \cdot 8 \\ 2 \cdot 7 + 5 \cdot 4 \\ 2 \cdot 0 + 5 \cdot 2 \end{pmatrix} = \begin{pmatrix} 54 & 34 & 10 \end{pmatrix}$$

2. Calculate the inverse of the orthogonal matrix

$$\boldsymbol{A} = \frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix}.$$

and check whether your result is correct. Also find the solution of $\boldsymbol{Ax} = \boldsymbol{b}$ with

$$\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

using the inverse matrix.

Solution:

The inverse of an orthogonal matrix is equal to its transposed. Thus,

$$\boldsymbol{A}^{-1} = \boldsymbol{A}^T = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{pmatrix}.$$

To check the result, the definition of the inverse matrix can be used:

$$\boldsymbol{A}^{-1}\boldsymbol{A} = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{I}.$$

Now the solution $\boldsymbol{x}$ can be found as

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \frac{1}{3} \begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}.$$

3. Calculate the L^p norm of the vector $\boldsymbol{y} = \begin{pmatrix} 4 & -3 \end{pmatrix}^T$ for $p = 1, 2, 3$.

Solution:

The definition of the L^p norm is

$$||\boldsymbol{y}||_p = \left(\sum_i |y_i|^p \right)^{\frac{1}{p}}.$$

Thus,

$$||\boldsymbol{y}||_1 = |y_1| + |y_2| = |4| + |-3| = 7,$$
$$||\boldsymbol{y}||_2 = \sqrt{|y_1|^2 + |y_2|^2} = \sqrt{|4|^2 + |-3|^2} = 5,$$

and

$$||\boldsymbol{y}||_3 = \sqrt[3]{|y_1|^3 + |y_2|^3} = \sqrt[3]{|4|^3 + |-3|^3} = \sqrt[3]{91} = 4.49794.$$

4. The following overdetermined, linear system of equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ is given, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{b} = \begin{pmatrix} 1.8 \\ 3.3 \\ 4.1 \end{pmatrix}.$$

Since there are more linear independent equations than unknowns, there is no solution. For this reason, the most reasonable solution should be found, by using the Moore-Penrose pseudoinverse

$$\boldsymbol{A}^+ = \boldsymbol{V}\boldsymbol{D}^+\boldsymbol{U}^T,$$

for which a singular value decomposition must be carried out. To perform a singular value decomposition, follow the steps below.

(a) Determine the eigenvalues λ_1 and λ_2 of $\boldsymbol{B} = \boldsymbol{A}^T\boldsymbol{A}$, where $\lambda_1 > \lambda_2$.

(b) Determine the Matrix D as $D = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \\ 0 & 0 \end{pmatrix}$

(c) Determine matrix $\boldsymbol{V} = \begin{pmatrix} \boldsymbol{v}^{(1)} & \boldsymbol{v}^{(2)} \end{pmatrix}$, where $\boldsymbol{v}^{(i)}$ are the eigenvectors of $\boldsymbol{B} = \boldsymbol{A}^T\boldsymbol{A}$. Note that $\boldsymbol{v}^{(1)}$ is the eigenvector to the larger eigenvalue λ_1.

(d) Determine matrix $\boldsymbol{U} = \begin{pmatrix} \boldsymbol{u}^{(1)} & \boldsymbol{u}^{(2)} & \boldsymbol{u}^{(3)} \end{pmatrix}$, where $\boldsymbol{u}^{(i)}$ are the eigenvectors of $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{A}^T$. Note that $\boldsymbol{u}^{(1)}$ is the eigenvector to the largest eigenvalue and $\boldsymbol{u}^{(3)}$ is the eigenvector to the smallest eigenvalue.

(e) Check weather $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices and weather $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$

(f) Determine the Matrix D^+.

(g) Determine the Moore-Penrose pseudoinverse A^+.

(h) Determine the most reasonable solution $\boldsymbol{x}$.

Hint:
Use the calculator or python to determine the eigenvalues and eigenvectors.

Solution:

(a) Matrix $\boldsymbol{B}$ is

$$\boldsymbol{B} = \begin{pmatrix} 3 & 12 \\ 12 & 56 \end{pmatrix}$$

and its eigenvalues are

$$\lambda_1 = 58.5904 \quad \text{and} \quad \lambda_2 = 0.4096.$$

(b) Matrix $\boldsymbol{D}$ is

$$\boldsymbol{D} = \begin{pmatrix} \sqrt{58.5940} & 0 \\ 0 & \sqrt{0.4096} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 7.6544 & 0 \\ 0 & 0.6400 \\ 0 & 0 \end{pmatrix}$$

(c) The eigenvectors of $\boldsymbol{B}$ are

$$\boldsymbol{v}^{(1)} = \begin{pmatrix} -0.2110 \\ -0.9775 \end{pmatrix} \quad \text{and} \quad \boldsymbol{v}^{(2)} = \begin{pmatrix} -0.9775 \\ 0.2110 \end{pmatrix},$$

whereby

$$\boldsymbol{V} = \begin{pmatrix} -0.2110 & -0.9775 \\ -0.9775 & 0.2110 \end{pmatrix}.$$

(d) Matrix $\boldsymbol{C}$ is

$$\boldsymbol{C} = \begin{pmatrix} 5 & 9 & 13 \\ 9 & 17 & 25 \\ 13 & 25 & 37 \end{pmatrix},$$

its eigenvalues are

$$\lambda_1 = 58.5904, \quad \lambda_2 = 0.4096, \quad \lambda_3 = 0,$$

and its eigenvectors are

$$\boldsymbol{u}^{(1)} = \begin{pmatrix} -0.2830 \\ -0.5384 \\ -0.7938 \end{pmatrix}, \quad \boldsymbol{u}^{(2)} = \begin{pmatrix} -0.8679 \\ -0.2085 \\ 0.4508 \end{pmatrix}, \quad \boldsymbol{u}^{(3)} = \begin{pmatrix} 0.4082 \\ -0.8165 \\ 0.4082 \end{pmatrix}.$$

Thus

$$\boldsymbol{U} = \begin{pmatrix} -0.2830 & -0.8679 & 0.4082 \\ -0.5384 & -0.2085 & -0.8165 \\ -0.7938 & 0.4508 & 0.4082 \end{pmatrix}.$$

(e) Orthogonality-check of $\boldsymbol{U}$ and $\boldsymbol{V}$:

$$\boldsymbol{U}\boldsymbol{U}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{I},$$

$$\boldsymbol{V}\boldsymbol{V}^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}.$$

Check of the singular value decomposition:

$$\boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T = \begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{pmatrix} = \boldsymbol{A}.$$

(f) Matrix $\boldsymbol{D}^+$ is

$$D^+ = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \\ 0 & 0 \end{pmatrix}^T = \begin{pmatrix} 0.1306 & 0 & 0 \\ 0 & 1.5625 & 0 \end{pmatrix}$$

(g) The Moore-Penrose pseudoinverse is

$$\boldsymbol{A}^+ = \boldsymbol{V}\boldsymbol{D}^+\boldsymbol{U}^T = \begin{pmatrix} 4/3 & 1/3 & -2/3 \\ -1/4 & 0 & 1/4 \end{pmatrix}.$$

(h) The most reasonable solution (least-square) is

$$\boldsymbol{x} = \boldsymbol{A}^+\boldsymbol{b} = \begin{pmatrix} 4/3 & 1/3 & -2/3 \\ -1/4 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1.8 \\ 3.3 \\ 4.1 \end{pmatrix} = \begin{pmatrix} 0.7667 \\ 0.5750 \end{pmatrix}.$$