Deep Learning for Speech and Language Processing
Exercise Sheet 1: Linear Algebra

Maximilian Schmidt, Pascal Tilli, Dirk Väth, Ngoc Thang Vu

November 2nd, 2020

Deadline

Please submit your solutions until Monday, November 9th, 2020, 5:30 p.m.. You may also see the deadline at any time on the website under *Exercises > Upload*.

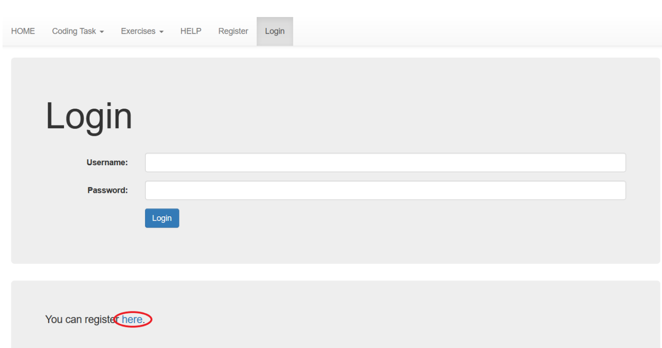
Notation

We try our best to be consistent with the mathematical notation throughout the course. Here are the most important conventions we use:

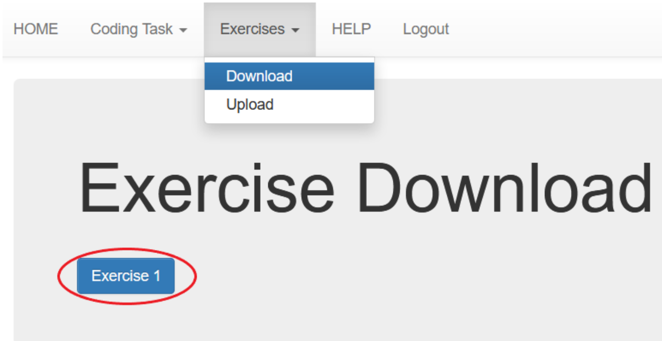
Symbol	Example	Description
greek letter	α	scalar
lower-case letter	x	vector (column vector)
...with one subscript	x_i	i-th entry of vector x (scalar)
...with two subscripts	x_{ij}	entry in the i-th row, j-th column of a matrix (scalar)
upper-case letter	M	matrix
	M^{-1}	inverse matrix of matrix M
	$v^\top,\ M^\top$	transposition of a vector or matrix
$\cdot$	$x \cdot y$	dot product of two vectors
superscript	M^l, b^l	weight matrix and bias at layer l of a neural network

Excercise Download and Submission

To obtain your personal excercise sheet and uploading your solutions, enter the following address in your web browser: https://dlcourse.ims.uni-stuttgart.de/ It will bring you to the login interface of the course page. If you don't have an account there yet, please register by clicking on the *register here* link.



You need to choose a username of your liking (please don't use anything offensive or inappropriate, otherwise you might be banned) and a password. Please provide your matriculation number, too. After confirming your registration, you will be able to log in and given an overview of your current excercise performance. Clicking on the *Excercises* section will offer you a choice to download your excercise sheet or to upload your solutions.



Troubleshooting: In case you forgot your password or you're having any trouble, please see the *HELP* section of the website.

Calculation Tasks

Submission Format

Upload (only solutions!) to https://dlcourse.ims.uni-stuttgart.de/ using your account as created in the previous section. **Please follow the required submission format precisely since evaluation is automatic (with a wrong format you might not get any points):**

- You can enter either single numbers or matrices.
- In case the result is incalculable, input *none*.
- Numbers can be entered as usual, e.g. 5.2 or as fractions, e.g. $-1/3$ (Note: don't include whitespace between sign and number).
- If you have a decimal number, round to 3 decimal places.
- Please make explicit all multiplication operations, e.g. always write $3 * x$ instead of $3x$.
- You may enter power expressions like $x\hat{}2$ or, equivalently, $x * *2$ instead of $x * x$ (**but not in case of the exponential function - use** $exp(x)$ **instead!**).
- Subscripts like x_1 can be entered as $x1$.
- Common functions such as exponential, logarithm and trigonometric functions may be entered as $ln(x), sin(x)$ etc.
- To enter a Matrix, type the elements of a row seperatetd by commas. Seperate rows by a semicolon (plus optional newline).

As an example, a 2×3 Matrix submission might look like

$$\begin{array}{ccc} -1, & 2 * x + y, & exp(3.2); \\ 4 * log(5 * x), & 5/2, & 6 * y * *2; \end{array}$$

. A 3-dimensional column vector will look like this: $1; 2; 3;$ and the corresponding row vector like this: $1, 2, 3;$

Warm-Up

Exercise 1.
Compute.

(1) $y = \alpha x + b$ for $\alpha = \frac{1}{7}, x = \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}, b = \begin{bmatrix} -8 \\ -5 \\ 6 \end{bmatrix}$

(2) $z = x \cdot y$ for $x = \begin{bmatrix} 2 \\ -10 \\ 10 \end{bmatrix}, y = \begin{bmatrix} -2 \\ -9 \\ 10 \end{bmatrix}$

(3) $C = AB$ for $A = \begin{bmatrix} -9 & 9 \\ 2 & -1 \end{bmatrix}, B = \begin{bmatrix} -2 & 5 \\ 10 & -6 \\ -10 & 2 \end{bmatrix}$

(4) $F = DE$ for $D = I_3, E = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

Derivatives

Exercise 2.
Compute the derivative wrt $x \in \mathbb{R}$ for the following functions:

(1) $f(x) = (7x + 6)^2$

(2) $g(x, y) = -2x^2 - 3x - 8y$

(3) $h(x) = \begin{bmatrix} 1x^2 \ln(1x) \\ -1x \end{bmatrix}$

Exercise 3.
Compute the derivative wrt $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ for the following functions:

(1) $f(x_1) = 8x_1 - 2$

(2) $g(x_1, x_2) = \begin{bmatrix} e^{-1x_1 - 9} \\ \frac{1}{6} e^{-4x_2} \\ 4x_1 2x_2 \end{bmatrix}$

Theory Tasks

Submission Format

You don't have to submit this part, but it will help you with the contents of this course.

Matrix Product

Exercise 4.
Show that matrix multiplication is not commutative.

Derivatives

Exercise 5.
Compute the derivative wrt $x \in \mathbb{R}^n$ for the following functions:

(1) $f(x) = x$

(2) $g(x) = x^\top \cdot x$ Hint: First compute $x^\top \cdot x$.

Linear Maps

Exercise 6.
The affine transformation $f(x) = Ax + b$ for matrix $A \in \mathbb{R}^{m \times n}$ input vector $x \in \mathbb{R}^n$ and bias $b \in \mathbb{R}^m$ is frequently referred to as 'linear activation' when used in neural networks, yet f is not a linear map by the definition of linear algebra. However, f can be rewritten as linear map by a small modification on the input x.

(1) Show that f is generally not a linear map.

(2) *(optional:)* Show how f can be rewritten as linear map. Hint: You need to add an additional dimension to the input x, such that $x' \in \mathbb{R}^{n+1}$.

(3) Deep neural networks consist of several layers, where each layer obtains the outputs of the previous layer and computes an activation function on them. Show that stacking several layers, where each layer has a linear activation function, does not improve the expressive power of the network over a single layer.