Introduction to Deep Learning for Speech and Language Processing
Exercise Sheet 2: Math for Machine Learning

28th October 2025

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
	I_n	n x n identity matrix
$\cdot$	$x \cdot y$	dot product of two vectors
superscript	W^l, b^l	weight matrix and bias at layer l of a neural network
∂	$\frac{\partial y}{\partial x}$	partial derivative of y with respect to x

For derivatives we use the numerator layout:

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix} \text{for } x \in \mathbb{R}^n, y \in \mathbb{R}^m$$

Linear Algebra

Linear Independence

Exercise 1.

Given are the vectors $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(1) Are the vectors u and v linearly independent?

(2) Is the set of all three vectors u, v, w linearly independent?

Solution 1.

(1) Yes, because the second component of v is always zero, regardless of scaling factor

(2) No, because e.g. $v + w = u$

Determinants and Eigenvalues

Exercise 2.

Given the matrix $A = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix}$ Compute.

(1) $A - \lambda I_2$

(2) $det(A - \lambda I_2)$

(3) Both eigenvalues of A

Solution 2.

- $\begin{bmatrix} 2-\lambda & 4 \\ 0 & 3-\lambda \end{bmatrix}$
- $det(A - \lambda I_2) = (2-\lambda)(3-\lambda)$
- $e_1 = 2$, $e_2 = 3$

Derivatives

Exercise 3.
Compute the derivative w.r.t. $x \in \mathbb{R}$ for the following functions:

(1) $f(x) = (-x-7)^2$

(2) $g(x, y) = -8x^2 + 5x + 7y$

(3) $h(x) = \begin{bmatrix} 2x^2 \ln(6x) \\ 3x \end{bmatrix}$

Solution 3.

(1) $\frac{d}{dx} f(x) = 2x + 14$

(2) $\frac{\delta}{\delta x} g(x, y) = 5 - 16x$

(3) $\frac{d}{dx} h(x) = \begin{bmatrix} 4x \ln(6x) + 2x^2 \cdot \frac{1}{6x} \cdot 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 4x \ln(6x) + 2x \\ 3 \end{bmatrix}$

Exercise 4.

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

(1) $f(x) = 9x_1 + 10x_2$

(2) $g(x_1, x_2) = \begin{bmatrix} e^{-4x_1+5} \\ \frac{1}{7} e^{7x_2} \\ 4x_1 9x_2 \end{bmatrix}$

Solution 4.

(1) $\nabla_x f(x) = \begin{bmatrix} 9 & 10 \end{bmatrix}$

(2) $J(g) = \begin{bmatrix} -4e^{-4x_1+5} & 0 \\ 0 & e^{7x_2} \\ 36x_2 & 36x_1 \end{bmatrix}$

Exercise 5.

Compute the derivative w.r.t. $x \in \mathbb{R}^n$ for the following functions:

(1) $f(x) = x$

Solution 5.

(1) $\frac{\delta f}{\delta x} = \begin{bmatrix} \frac{\delta x_1}{\delta x_1} & \cdots & \frac{\delta x_1}{\delta x_n} \\ \vdots & \ddots & \vdots \\ \frac{\delta x_n}{\delta x_1} & \cdots & \frac{\delta x_n}{\delta x_n} \end{bmatrix} = I_n$