Introduction to Deep Learning for Speech and Language Processing
Exercise Sheet 1: Math for Machine Learning

17th 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

Linear Algebra

Vector Operations

Exercise 1.
Compute.

(1) $y = \alpha x + b$ for $\alpha = \frac{1}{6}, x = \begin{bmatrix} -2 \\ 9 \\ 3 \\ -4 \end{bmatrix}, b = \begin{bmatrix} 1 \\ \frac{1}{2} \\ -\frac{1}{2} \\ -2 \end{bmatrix}$

(2) $z = x \cdot y = \langle x, y \rangle$ for $x = \begin{bmatrix} 1 \\ -2 \\ -\frac{1}{2} \\ 3 \end{bmatrix}, y = \begin{bmatrix} -3 \\ 7 \\ 2 \\ 2 \end{bmatrix}$

(3) $z = x \odot y$ for $x = \begin{bmatrix} 1 \\ 0 \\ 0.5 \end{bmatrix}, y = \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix}$

Solution 1.

(1) $y = \begin{bmatrix} \frac{2}{3} \\ 2 \\ 0 \\ -\frac{8}{3} \end{bmatrix}$

(2) $z = -12$

(3) $z = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}$

Vector Similarity

Exercise 2.

The following vectors represent word embeddings: $dog = \begin{bmatrix} 2.0 \\ 1.5 \\ 3.0 \end{bmatrix}$, $puppy = \begin{bmatrix} 3.1 \\ 1.25 \\ 4 \end{bmatrix}$, $car = \begin{bmatrix} -4.7 \\ 2.1 \\ -5.5 \end{bmatrix}$.

Compute the pair-wise cosine similarity between

(1) *dog* and *puppy*

(2) *dog* and *car*

.

Solution 2.

(1) $cos(\theta) = 0.9861\ldots$

(2) $cos(\theta) = -0.77333$

Linear Independence

Exercise 3.

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 3.

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

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

Matrix Operations

Exercise 4.

Compute.

(1) A^T for $A = \begin{bmatrix} 3 & 1 & 0 \\ -1 & -2 & 5 \end{bmatrix}$

(2) $Wx + b$ for $W = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -2 & \frac{1}{2} \end{bmatrix}$, $x = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(3) $C = AB$ and $D = BA$ for $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -2 & \frac{1}{2} \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 \\ 0 & 1 \\ -2 & -1 \end{bmatrix}$

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

Solution 4.

(1) $A^T = \begin{bmatrix} 3 & -1 \\ 1 & -2 \\ 0 & 5 \end{bmatrix}$

(2) $Wx + b = \begin{bmatrix} 3 \\ 4.5 \end{bmatrix}$

(3) $C = \begin{bmatrix} 4 & 6 \\ 3 & 3.5 \end{bmatrix}$,

$D = \begin{bmatrix} 8 & -2 & -\frac{1}{2} \\ 2 & -2 & \frac{1}{2} \\ -4 & -2 & \frac{3}{2} \end{bmatrix}$

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