

tl;dr: Linear algebra is the branch of mathematics concerning matrices and vectors. The most important operation is matrix multiplication. Please get very comfortable with it.

## Linear Algebra

Linear algebra is the branch of mathematics concerning systems of linear equations, the study of vectors, vector spaces, linear transformations, and matrices. In contrast to calculus, which is marked by bends and twists on curved surfaces, linear algebra is a study of straight lines and planes—be it higher-dimensional. Perhaps the most important concept of linear algebra is that of the matrix and the corresponding operations (addition, subtraction, and multiplication).

### Notes

#### Numbers

“God made the natural numbers; all else is the work of man.” — Leopold Kronecker (maybe). The natural numbers  $\mathbb{N} = 1, 2, 3, 4, 5, \dots$  along with the basic operations of addition and multiplication and limits (calculus) and roots, can be used to construct the rational numbers  $\mathbb{Q} = p/q | p, q \in \mathbb{N}$ , the real numbers  $\mathbb{R}$ , and even the complex numbers  $\mathbb{C}$  (which will not be featured in this course). The algebraic completeness theorem states that we cannot escape the complex numbers using the aforementioned operations. The complex numbers are the end of the line. And yet, much of the world’s phenomena seems inadequately described by a single number.

#### Vectors

A vector is an array of numbers, thought of as a single object. Similar to numbers, vectors have associated operations of addition and scalar multiplication. Vectors can be added together to form new vectors, and they can be multiplied by scalars. They can be thought of as representing points in space, displacements, velocities, forces, features of data, or solutions to systems of linear equations. In this course, we will be working with  $n$  dimensional vectors  $v \in \mathbb{R}^n$ .

**Dot Product** The dot product (also called inner product) of two vectors  $v, w \in \mathbb{R}^n$  is a scalar defined as:

$$v \cdot w = \sum_{i=1}^n v_i w_i$$

#### Matrices

A matrix is a rectangular array of numbers (or a flat array of vectors), thought of as a single object. Matrices too have associated operations of addition and scalar multiplication<sup>1</sup>, and most importantly, matrix multiplication. Matrix multiplication is a generalization of the dot

---

<sup>1</sup>and the trace, the determinants, and more (which we will not use)

product of vectors, and it is the most important operation in linear algebra, and the most common operation in scientific computing (I've been told).

Matrices are *the* way to represent linear transformations, images, films, music, machine learning models—basically all kinds of data in the world.

**Matrix Multiplication** Matrix multiplication is a generalization of the dot product of vectors. It is defined as follows: Let  $A \in \mathbb{R}^{m \times n}$  and  $B \in \mathbb{R}^{n \times p}$ . Then the product  $C = AB \in \mathbb{R}^{m \times p}$  is defined as follows:

$$C_{ij} = \sum_{k=1}^n A_{ik} B_{kj}$$

## JAX

We will be using JAX, a numerical computing library, to work with vectors and matrices. JAX is best used in conjunction with pen and paper, for understanding and debugging.

## Exercises

### Pen and Paper

1. Compute the dot product of the vectors  $v = [1, 2, 3]$  and  $w = [4, 5, 6]$ .
2. Compute the product  $AB$  of the matrices  $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ .
3. Compute the product  $AB$  of the matrices  $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ .
4. Compute the product  $AB$  of the matrices  $A = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ .
5. Compute the product of  $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$ .
6. Compute the product of  $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ .
7. Philosophize on the difference between the results of exercises 5 and 6.
8. Philosophize on exercise 3. Set  $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  and compute  $AB$ .

## Code

I use zed.dev as my editor, and recommend you do the same for this course. The repo of the course includes a `requirements.txt` file with the necessary dependencies, but can be installed with `pip install aigs` in a clean virtual environment. This should give you all necessary dependencies for the course.

1. Ensure you have a working environment with the course dependencies.

```
import jax.numpy as jnp # JAX's version of numpy
from jax import random, tree, vmap, grad, nn # JAX's utilities
import optax # library for optimization
import chex # library for checking JAX types (nice for mental health)
from typing import List # for type hints
```

- Define some type aliases for convenience.

```

Number = float | int  # useful for type hints
Vector = List[Number]  # useful for type hints
Matrix = List[Vector]  # useful for type hints

```

2. Implement the dot product of two vectors in python without JAX.

```

def dot(v: Vector, w: Vector) -> Number:
    pass
    # return sum(v_i * w_i for v_i, w_i in zip(v, w))

```

3. Implement the matrix multiplication of two matrices in python without JAX (but with your previous function)

```

def matmul(A: Matrix, B: Matrix) -> Matrix:
    pass
    # return [[dot(row, col) for col in zip(*B)] for row in A]

```

4. Review `1a.py`
5. Tweak values of `lr`, and see what happens.
6. Put `params` into a chex `@dataclass`, and discuss why this is a good idea.
7. Write a function that counts the number of parameters in a `params`. hint: use `tree`.