From Math to Machines: A Comprehensive Guide to ML and DL

Covering Calculus, Linear Algebra and Statistics

Francesco Danese

PUBLISHED BY NOBODY

BOOK-WEBSITE.COM

First printing, March 2022


Our company
Receipts
Sales
Orders
Business items
Sales
Orders

Contents

List of Figures

List of Tables

Linear Algebra

1. Vectors and Matrices

1.1 Vectors

In the vast field of mathematics, vectors play a fundamental role in various branches from physics and engineering to computer science and data science. A vector can be thought as a ordered sequence of numbers that defines a directed line segment in space, an arrow pointing from one point to another, capturing both the **magnitude** (its length) and the **direction** in which it is pointing. We can denote a vector using a bold letter such as " $\boldsymbol{v}$ ", and writing down its components, for example:

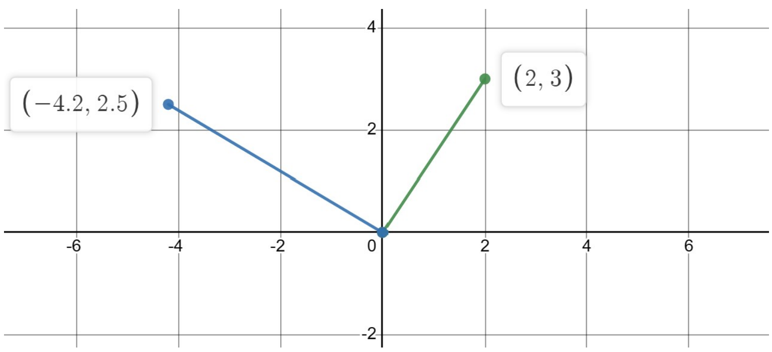
$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \mathbf{w} = \begin{bmatrix} -4.2 \\ 2.5 \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} 2.1 \\ 0 \\ -3 \end{bmatrix}$$

Each vector has a **dimensionality** (i.e.the number of its components) that tells the space in which it lives, denoted as R^D. Vectors **v** and **w** above live in R^2 while **u** lives in R^3. The representation is usually column-wise, but we can obtain the row-wise version of a vector by a **transposition** operation denoted as $\mathbf{v}^T$, for instance:

$$\mathbf{v}^T = \begin{bmatrix} 2 & 3 \end{bmatrix} \qquad \mathbf{w}^T = \begin{bmatrix} -4.2 & 2.5 \end{bmatrix} \qquad \mathbf{u}^T = \begin{bmatrix} 2.1 & 0 & -3 \end{bmatrix}$$

We can visualize a 2-dimensional vector by plotting it into a cartesian plane, where the two components are the shifts in the x and y directions respectively:

Figure 1.1: Vectors v and w graphically.

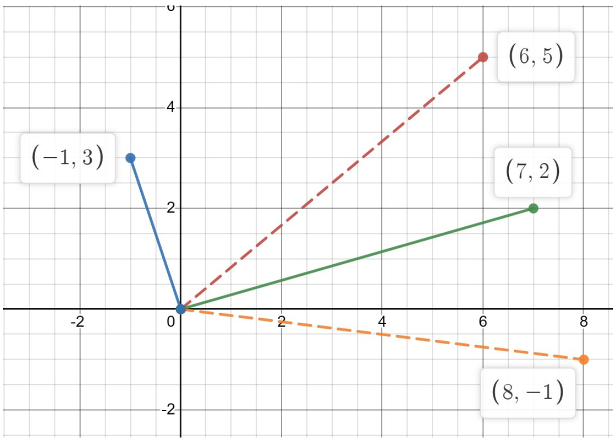


1.1.1 Addition and Subtraction

Vector addition is a fundamental operation that combines two vectors to create a new vector. When adding vectors, we place the tail of each subsequent vector at the head of the previous vector. The resultant vector is then drawn from the tail of the first vector (usually the origin) to the head of the final vector. The sum of two vectors represents the combined effect of their individual magnitudes and directions, and is calculated by adding their corresponding components. If $\mathbf{u} = [u_1, u_2, u_3]$ and $\mathbf{v} = [v_1, v_2, v_3]$ then $\mathbf{u}+\mathbf{v} = [u_1+v_1, u_2+v_2, u_3+v_3]$. For what regards the vector subtraction we have that: $\mathbf{v}-\mathbf{u} = \mathbf{v}+(-\mathbf{u})$. For example:

$$\begin{bmatrix}7\\2\end{bmatrix}+\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}6\\5\end{bmatrix} \qquad \begin{bmatrix}7\\2\end{bmatrix}-\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}8\\-1\end{bmatrix}$$

Figure 1.2: Addition (red) and subtraction (orange) of the blue and green vectors.

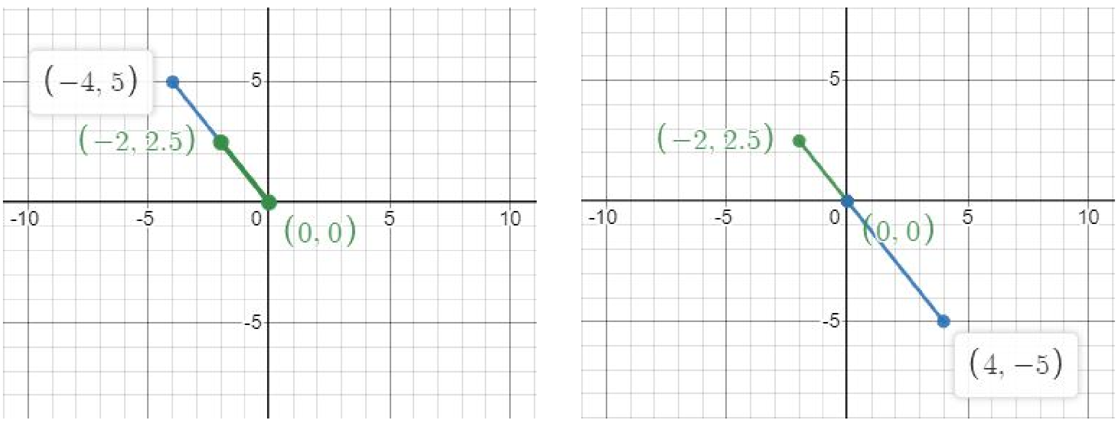


1.1.2 Scalar multiplication

Scalar multiplication involves scaling a vector by a scalar, which is a real number. The vector's magnitude is multiplied by the scalar value, while its direction remains unchanged (or flipped if the scalar is negative). Mathematically, scalar multiplication is performed by multiplying each component of the vector by the scalar. If $\mathbf{v} = [v_1, v_2, v_3]$ and c is a scalar, then $c\mathbf{v} = [cv_1, cv_2, cv_3]$.

$$2\cdot\begin{bmatrix}-2\\2.5\end{bmatrix}=\begin{bmatrix}-4\\5\end{bmatrix} \qquad -2\cdot\begin{bmatrix}-2\\2.5\end{bmatrix}=\begin{bmatrix}4\\-5\end{bmatrix}$$

Figure 1.3: Green vector scaled by a factor of 2 (left) and -2 (right)



1.1.3 Dot Product

The dot product, also called inner product, takes two vectors and produces a *scalar value*. It is denoted by a dot $(\cdot)$ and it's the sum of the products of the corresponding components of the two vectors. If $\mathbf{v} = [v_1, v_2, ..., v_n]$ and $\mathbf{w} = [w_1, w_2, ..., w_n]$ are two vectors in a n-dimensional space, then:

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

For example:

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ -1 \end{bmatrix} = 2 \cdot 4 + 3 \cdot (-1) = 8 - 3 = 5$$

We also some important geometric interpretation. The magnitude[1] of a vector $\mathbf{v}$ is always denoted as $||\mathbf{v}||$, and if we indicate the angle between the two vectors as θ we have another definition for the dot product:

$$\mathbf{v} \cdot \mathbf{w} = ||\mathbf{v}|| \cdot ||\mathbf{w}|| \cdot \cos(\theta)$$

And it follows that:

$$\cos(\theta) = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|| \cdot ||\mathbf{w}||} \qquad \textit{and} \qquad \theta = \arccos(\frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|| \cdot ||\mathbf{w}||})$$

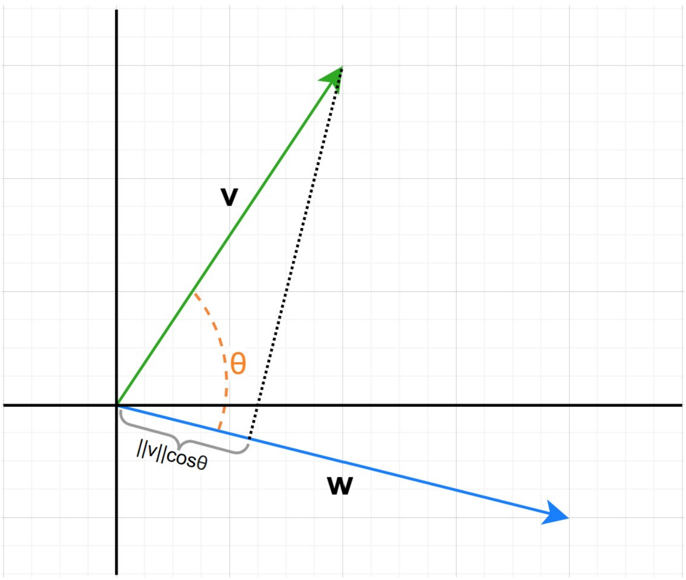


Figure 1.4: $||\mathbf{v}||\cos(\theta)$ is the projection $\mathbf{v}$ on $\mathbf{w}$

It comes as a conclusion that if $\mathbf{v}$ and $\mathbf{w}$ are unit vectors, meaning that their norm is equal to 1 ($||\mathbf{v}|| = ||\mathbf{w}|| = 1$), then the dot product represents exactly the cosine of the angle between them. The dot product is commutative and can be done only if the vectors have the same dimension.

1.1.4 Norms

Informally, the norm of a vector represents how "big" is the vector. It is a function $||\cdot||$ that maps $\mathbf{v}$ to a scalar. For any norm we have the following properties:

$$||c\mathbf{v}|| = |c| \cdot ||\mathbf{v}||, \qquad ||\mathbf{v} + \mathbf{w}|| = ||\mathbf{v}|| + ||\mathbf{w}||, \qquad ||\mathbf{v}|| > 0 \quad \forall \mathbf{v} \neq \mathbf{0}$$

[1] The magnitude is the l_2-norm: see next section about vector norms

Where c is a scalar value. Different norms encode different notions of size. The generalized formula, called ℓ_p norm is:

$$||\mathbf{x}||_p = (\sum_{i=1}^{d} |x_i|^p)^{\frac{1}{p}}$$

We then have the ℓ_1 norm, also called the *Manhattan distance*:

$$||\mathbf{x}||_1 = \sum_{i=1}^{d} |x_i|$$

And the ℓ_2 norm, also called the *Euclidean Distance*:

$$||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$$

Dividing a vector $\mathbf{v}$ by its norm is a common operation called *normalizing*, that result in a unit vector pointing in the same direction of $\mathbf{v}$ (but with norm = 1):

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{||\mathbf{v}||}$$

1.2 Matrices

A matrix is a collection of n column vectors in M-dimension or m row vectors in N-dimension.

$$\begin{bmatrix} x_{11} & \dots & \dots & x_{1n} \\ \vdots & \ddots & & \dots \\ x_{m1} & \dots & \ddots & x_{mn} \end{bmatrix}$$

The identity matrix is a diagonal matrix where all elements in the diagonal are 1s and all the others are 0s:

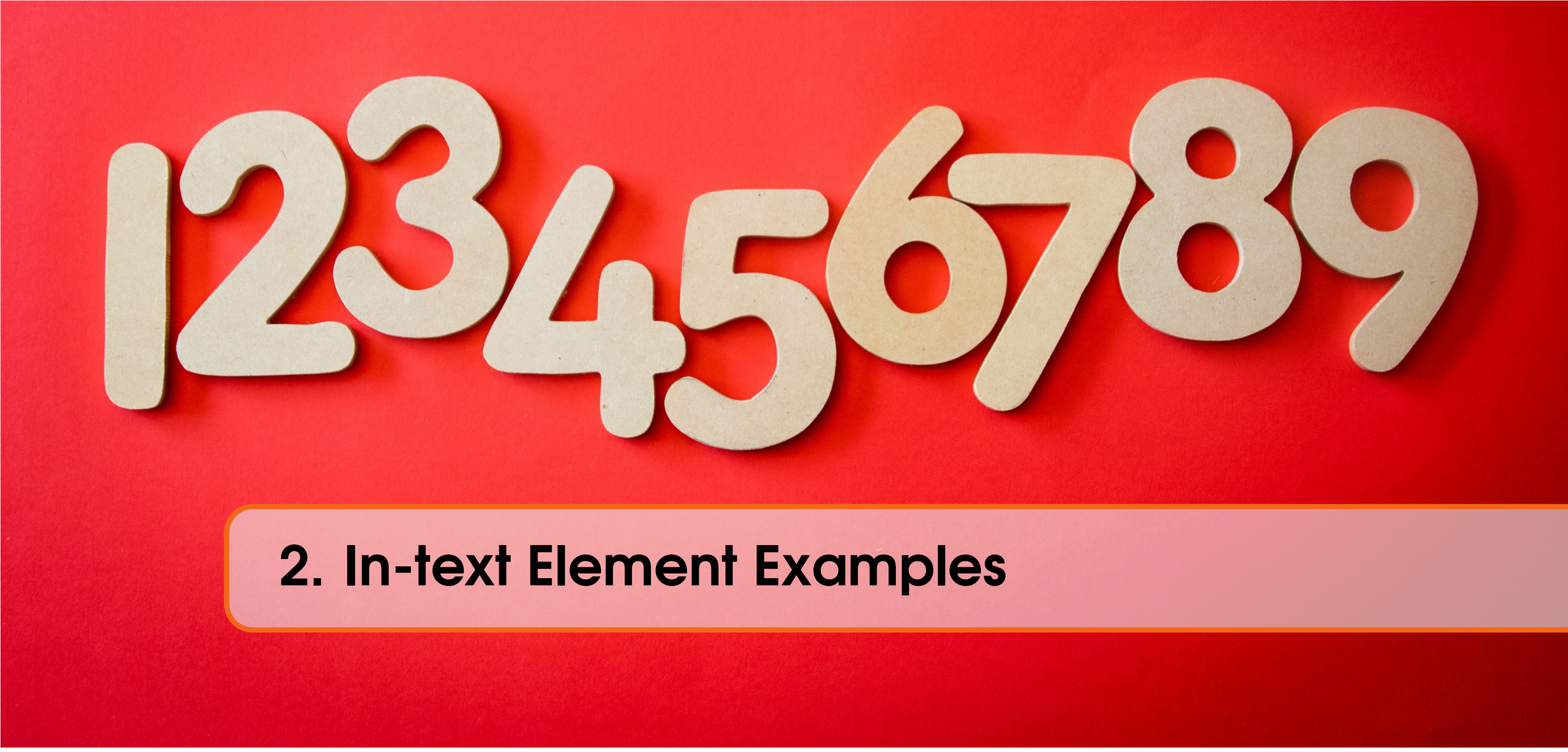
$$I_3 = diag(1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The trace of a matrix is the sum of the diagonal elements:

$$trace(A) = \sum_i A_{ii}$$

Transposing a matrix means swapping its rows with its columns; The first row becomes the first column, the second row becomes the second column and so on. We have the following properties:

$$(A^T)^T = A \qquad (AB)^T = A^T B^T \qquad (A+B)^T = A^T + B^T$$



2. In-text Element Examples

2.1 Referencing Publications

This statement requires citation [1]; this one is more specific [2, page 162].

2.2 Link Examples

This is a URL link: LaTeX Templates. This is an email link: example@example.com. This is a monospaced URL link: `https://www.LaTeXTemplates.com`.

2.3 Lists

Lists are useful to present information in a concise and/or ordered way.

2.3.1 Numbered List

1. First numbered item
    a. First indented numbered item
    b. Second indented numbered item
        i. First second-level indented numbered item
2. Second numbered item
3. Third numbered item

2.3.2 Bullet Point List

- First bullet point item
    - First indented bullet point item
    - Second indented bullet point item
        - First second-level indented bullet point item
- Second bullet point item
- Third bullet point item

2.3.3 Descriptions and Definitions

Name Description
Word Definition
Comment Elaboration

2.4 International Support

àáâäãåèéêëìíîïòóôöõøùúûüÿýñçčšž
ÀÁÂÄÃÅÈÉÊËÌÍÎÏÒÓÔÖÕØÙÚÛÜŸÝÑ
ßÇŒÆČŠŽ

2.5 Ligatures

fi fj fl ffl ffi Ty

II Part Two Title

3. Mathematics

3.1 Theorems

3.1.1 Several equations

This is a theorem consisting of several equations.

Theorem 3.1 — Name of the theorem. In $E = \mathbb{R}^n$ all norms are equivalent. It has the properties:

$$\big| ||\mathbf{x}|| - ||\mathbf{y}|| \big| \leq ||\mathbf{x} - \mathbf{y}|| \tag{3.1}$$

$$|| \sum_{i=1}^{n} \mathbf{x}_i || \leq \sum_{i=1}^{n} ||\mathbf{x}_i|| \quad \text{where } n \text{ is a finite integer} \tag{3.2}$$

3.1.2 Single Line

This is a theorem consisting of just one line.

Theorem 3.2 A set $\mathscr{D}(G)$ in dense in $L^2(G)$, $|\cdot|_0$.

3.2 Definitions

A definition can be mathematical or it could define a concept.

Definition 3.1 — Definition name. Given a vector space E, a norm on E is an application, denoted $||\cdot||$, E in $\mathbb{R}^+ = [0, +\infty[$ such that:

$$||\mathbf{x}|| = 0 \Rightarrow \mathbf{x} = \mathbf{0} \tag{3.3}$$

$$||\lambda \mathbf{x}|| = |\lambda| \cdot ||\mathbf{x}|| \tag{3.4}$$

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}|| \tag{3.5}$$

3.3 Notations

■ **Notation 3.1** Given an open subset G of $\mathbb{R}^n$, the set of functions φ are:

1. Bounded support G;
2. Infinitely differentiable;

a vector space is denoted by $\mathscr{D}(G)$.

3.4 Remarks

This is an example of a remark.

R The concepts presented here are now in conventional employment in mathematics. Vector spaces are taken over the field $\mathbb{K}=\mathbb{R}$, however, established properties are easily extended to $\mathbb{K}=\mathbb{C}$.

3.5 Corollaries

Corollary 3.1 — Corollary name. The concepts presented here are now in conventional employment in mathematics. Vector spaces are taken over the field $\mathbb{K}=\mathbb{R}$, however, established properties are easily extended to $\mathbb{K}=\mathbb{C}$.

3.6 Propositions

3.6.1 Several equations

Proposition 3.1 — Proposition name. It has the properties:

$$\left| ||\mathbf{x}|| - ||\mathbf{y}|| \right| \leq ||\mathbf{x}-\mathbf{y}|| \tag{3.6}$$

$$||\sum_{i=1}^{n} \mathbf{x}_i|| \leq \sum_{i=1}^{n} ||\mathbf{x}_i|| \quad \text{where } n \text{ is a finite integer} \tag{3.7}$$

3.6.2 Single Line

Proposition 3.2 Let $f,g \in L^2(G)$; if $\forall \varphi \in \mathscr{D}(G)$, $(f,\varphi)_0 = (g,\varphi)_0$ then $f=g$.

3.7 Examples

3.7.1 Equation Example

■ **Example 3.1** Let $G=\{x\in\mathbb{R}^2 : |x|<3\}$ and denoted by: $x^0=(1,1)$; consider the function:

$$f(x)=\begin{cases} e^{|x|} & \text{si } |x-x^0| \leq 1/2 \\ 0 & \text{si } |x-x^0| > 1/2 \end{cases} \tag{3.8}$$

The function f has bounded support, we can take $A=\{x\in\mathbb{R}^2 : |x-x^0|\leq 1/2+\varepsilon\}$ for all $\varepsilon \in]0;5/2-\sqrt{2}[$. ■

3.7.2 Text Example

■ **Example 3.2 — Example name.** Aliquam arcu turpis, ultrices sed luctus ac, vehicula id metus. Morbi eu feugiat velit, et tempus augue. Proin ac mattis tortor. Donec tincidunt, ante rhoncus luctus semper, arcu lorem lobortis justo, nec convallis ante quam quis lectus. Aenean tincidunt sodales massa, et hendrerit tellus mattis ac. Sed non pretium nibh. Donec cursus maximus luctus. Vivamus lobortis eros et massa porta porttitor. ■

3.8 Exercises

Exercise 3.1 This is a good place to ask a question to test learning progress or further cement ideas into students' minds. ■

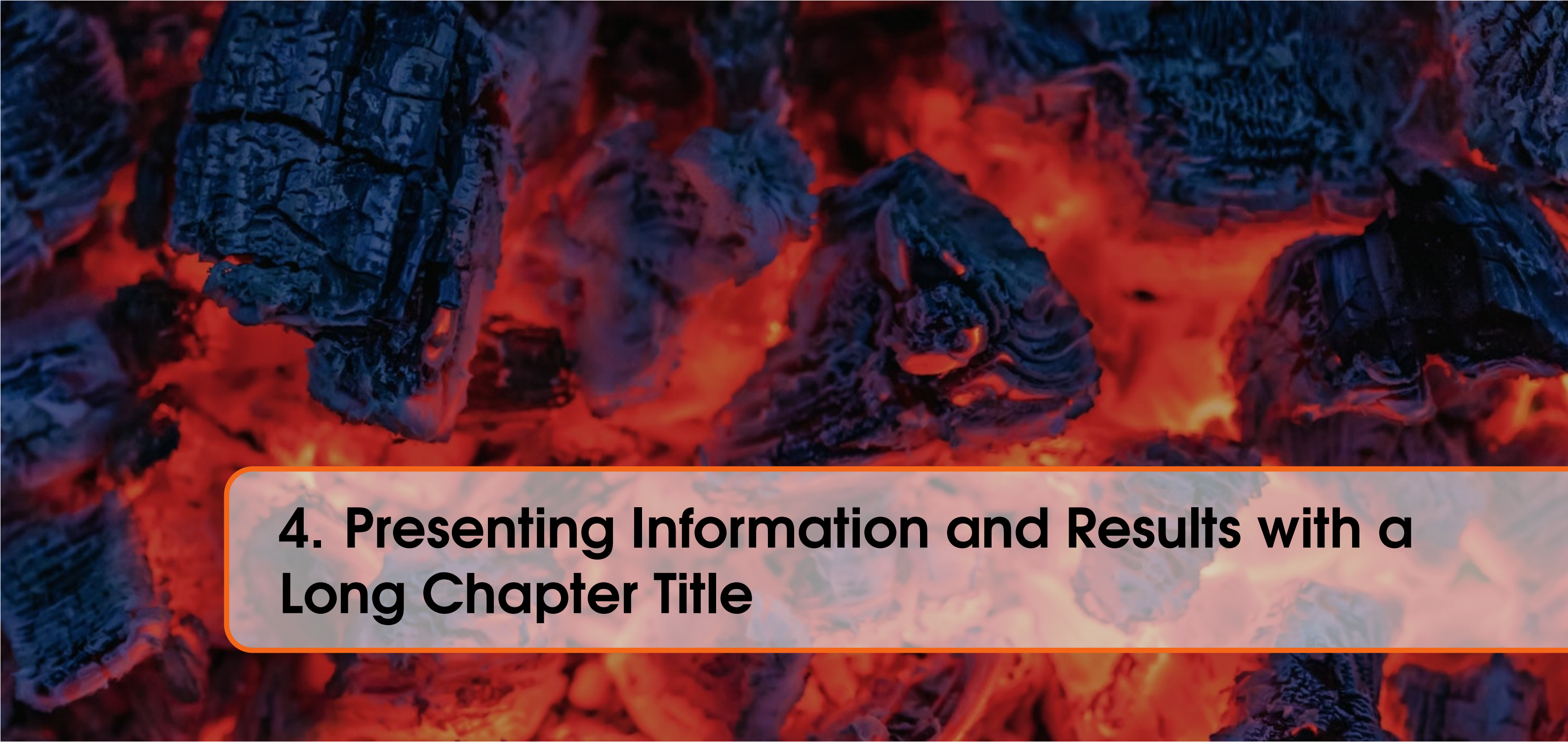
3.9 Problems

Problem 3.1 What is the average airspeed velocity of an unladen swallow?

3.10 Vocabulary

Define a word to improve a students' vocabulary.

- **Vocabulary 3.1 — Word.** Definition of word.



4. Presenting Information and Results with a Long Chapter Title

4.1 Table

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent porttitor arcu luctus, imperdiet urna iaculis, mattis eros. Pellentesque iaculis odio vel nisl ullamcorper, nec faucibus ipsum molestie. Sed dictum nisl non aliquet porttitor. Etiam vulputate arcu dignissim, finibus sem et, viverra nisl. Aenean luctus congue massa, ut laoreet metus ornare in. Nunc fermentum nisi imperdiet lectus tincidunt vestibulum at ac elit. Nulla mattis nisl eu malesuada suscipit.

Treatments	**Response 1**	**Response 2**
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 4.1: Table caption.

Referencing Table 4.1 in-text using its label.

4.2 Figure

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Praesent porttitor arcu luctus, imperdiet urna iaculis, mattis eros. Pellentesque iaculis odio vel nisl ullamcorper, nec faucibus ipsum molestie. Sed dictum nisl non aliquet porttitor. Etiam vulputate arcu dignissim, finibus sem et, viverra nisl. Aenean luctus congue massa, ut laoreet metus ornare in. Nunc fermentum nisi imperdiet lectus tincidunt vestibulum at ac elit. Nulla mattis nisl eu malesuada suscipit.

Figure 4.1: Figure caption.

Referencing Figure 4.1 in-text using its label.

Treatments	**Response 1**	**Response 2**
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 4.2: Floating table.

Figure 4.2: Floating figure.

Bibliography

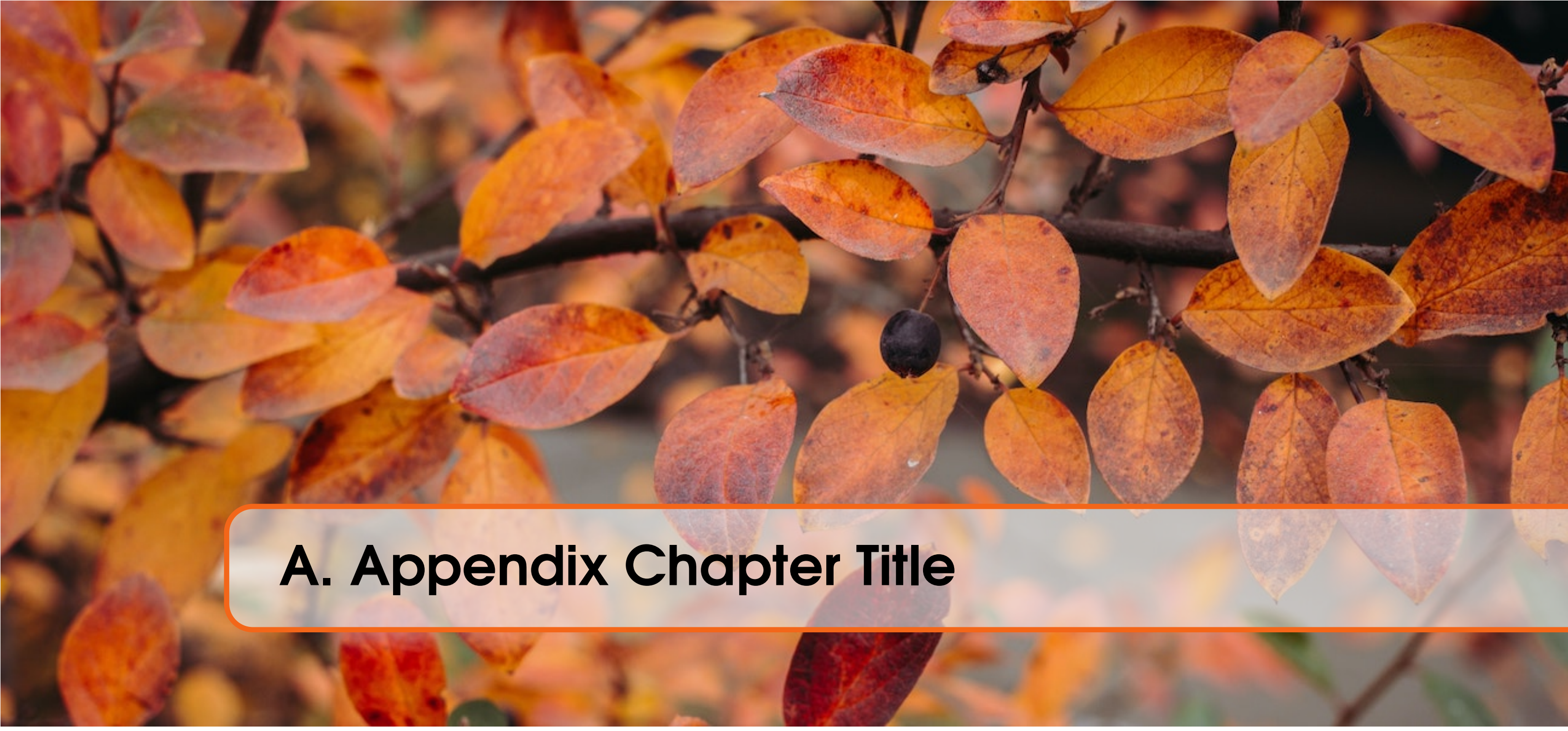
Articles

[1] A. B. Jones and J. M. Smith. “Article Title”. In: *Journal title* 13.52 (Mar. 2022), pages 123–456. DOI: 10.1038/s41586-021-03616-x (cited on page 15).

Books

[2] J. M. Smith and A. B. Jones. *Book Title*. 7th. Publisher, 2021 (cited on page 15).

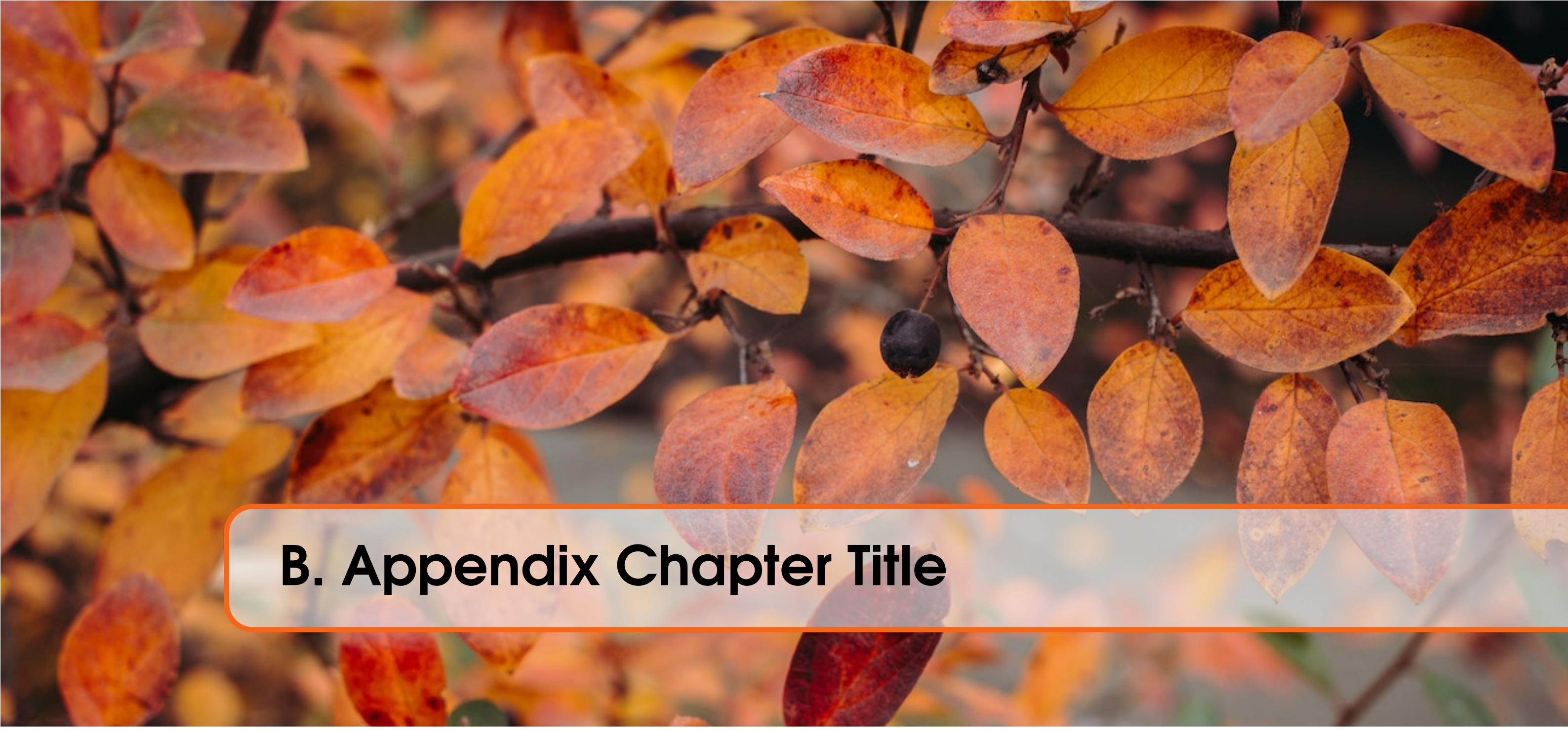
Index



A. Appendix Chapter Title

A.1 Appendix Section Title

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam auctor mi risus, quis tempor libero hendrerit at. Duis hendrerit placerat quam et semper. Nam ultricies metus vehicula arcu viverra, vel ullamcorper justo elementum. Pellentesque vel mi ac lectus cursus posuere et nec ex. Fusce quis mauris egestas lacus commodo venenatis. Ut at arcu lectus. Donec et urna nunc. Morbi eu nisl cursus sapien eleifend tincidunt quis quis est. Donec ut orci ex. Praesent ligula enim, ullamcorper non lorem a, ultrices volutpat dolor. Nullam at imperdiet urna. Pellentesque nec velit eget est euismod pretium.



B. Appendix Chapter Title

B.1 Appendix Section Title

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam auctor mi risus, quis tempor libero hendrerit at. Duis hendrerit placerat quam et semper. Nam ultricies metus vehicula arcu viverra, vel ullamcorper justo elementum. Pellentesque vel mi ac lectus cursus posuere et nec ex. Fusce quis mauris egestas lacus commodo venenatis. Ut at arcu lectus. Donec et urna nunc. Morbi eu nisl cursus sapien eleifend tincidunt quis quis est. Donec ut orci ex. Praesent ligula enim, ullamcorper non lorem a, ultrices volutpat dolor. Nullam at imperdiet urna. Pellentesque nec velit eget est euismod pretium.