Learning Representations with Physics Constraints

Master's Thesis

Maria Papadaki

Supervisor: Prof. A. Einstein

Dept. of Applied Mathematics

University of Crete

June 2025

Acknowledgments

Thanks to my cat.

To my dog.

Committee

Prof. A. Einstein, Chair
Prof. B. Heisenberg
Dr. C. Schrodinger

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Chapter 1

Package Demonstration

This chapter demonstrates the mathematical and typographical capabilities built into the thesis class.

1.1 Mathematical Notation

1.1.1 Basic Mathematics

The fundamental theorem of calculus states that if f is continuous on $[a, b]$ and F is an antiderivative of f, then:

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a) \tag{1.1}$$

1.1.2 Vectors and Bold Mathematics

Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ be a vector in $\mathbb{R}^n$. The gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ is denoted as $\vec{\nabla} f(\boldsymbol{x})$ or $\nabla f(\boldsymbol{x})$.

For machine learning, we often work with:

$$\begin{aligned} \boldsymbol{W} &\in \mathbb{R}^{m\times n} \quad \text{(weight matrix)} && (1.2)\\ \boldsymbol{b} &\in \mathbb{R}^{m} \quad \text{(bias vector)} && (1.3)\\ \boldsymbol{y} &= \boldsymbol{W}\boldsymbol{x} + \boldsymbol{b} \quad \text{(linear transformation)} && (1.4) \end{aligned}$$

1.1.3 Calculus and Differentials

The total differential of a function $f(x, y)$ is:

$$\mathrm{d}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y \tag{1.5}$$

For partial derivatives, we use:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \nabla^2 f \tag{1.6}$$

1.1.4 Optimization Operators

The thesis class includes predefined mathematical operators for optimization:

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \tag{1.7}$$

$$\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}\|^2 + \lambda\|\boldsymbol{w}\|^2 \tag{1.8}$$

These operators automatically handle proper spacing and formatting in optimization problems.

1.1.5 Probability and Statistics

The thesis class includes flexible commands for probability and statistics notation:

Basic Probability Notation

Working with probability spaces and distributions:

$$X \sim \mathcal{N}(\mu, \sigma^2) \quad \text{(normal distribution)} \tag{1.9}$$

$$Y \sim \mathcal{U}(a, b) \quad \text{(uniform distribution)} \tag{1.10}$$

$$Z \in \mathbb{R} \quad \text{(real-valued random variable)} \tag{1.11}$$

Expectation Operator

The `\Mean` command adapts to different contexts:

$$\mathbb{E} \quad \text{(plain expectation symbol)} \tag{1.12}$$

$$\mathbb{E}_X \quad \text{(expectation w.r.t. X)} \tag{1.13}$$

$$\mathbb{E}_X\left[Y\right] = \mathbb{E}_X\left[Y\right] \quad \text{(expectation of Y w.r.t. X)} \tag{1.14}$$

$$\mathbb{E}_\mu\left[X^2\right] = \mathbb{E}_\mu\left[X^2\right] \quad \text{(second moment)} \tag{1.15}$$

Probability Operator

The `\Prob` command works similarly:

$$\mathbb{P} \quad \text{(plain probability symbol)} \tag{1.16}$$

$$\mathbb{P}_X \quad \text{(probability w.r.t. X)} \tag{1.17}$$

$$\mathbb{P}_X\left(A\right) = \mathbb{P}_X\left(A\right) \quad \text{(probability of event A)} \tag{1.18}$$

$$\mathbb{P}_\mu\left(X > 0\right) = \mathbb{P}_\mu\left(X > 0\right) \quad \text{(conditional probability)} \tag{1.19}$$

Variance Operator

The `\Var` command for variance calculations:

$$\mathbb{V} \quad \text{(plain variance symbol)} \tag{1.20}$$

$$\mathbb{V}_X \quad \text{(variance w.r.t. X)} \tag{1.21}$$

$$\mathbb{V}_X\left[Y\right] = \mathbb{V}_X\left[Y\right] \quad \text{(variance of Y w.r.t. X)} \tag{1.22}$$

$$\mathbb{V}_\theta\left[X\right] = \mathbb{V}_\theta\left[X\right] \quad \text{(parametric variance)} \tag{1.23}$$

Combined Example

A typical machine learning probability statement:

$$\mathbb{E}_\theta\left[\ell(X,Y)\right] = \mathbb{E}_\theta\left[\mathbb{P}_{X,Y}\left(Y \neq f_\theta(X)\right)\right] \tag{1.24}$$

where ℓ is the loss function and f_θ is our model.

1.1.6 Special Sets and Greek Letters

We work with various mathematical sets:

- Natural numbers: $\mathbb{N} = \{1, 2, 3, \ldots\}$
- Real numbers: $\mathbb{R}$
- Complex numbers: $\mathbb{C}$
- Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

Greek letters in upright form: $\pi \approx 3.14159$, $e = 2.71828$.

1.1.7 Theorems and Proofs

Theorem 1.1 (Pythagorean Theorem)**.** *In a right triangle with legs of length a and b, and hypotenuse of length c, we have:*

$$a^2 + b^2 = c^2 \tag{1.25}$$

Proof. The proof is left as an exercise for the reader. □

1.2 Cross-References

This section demonstrates smart referencing with cleveref. We can reference Eq. (1.1), Theorem 1.1, and Chapter 1 automatically.

Multiple references work too: Eq. (1.1) and Theorem 1.1 show different mathematical concepts. Use `cref` in the middle of the sentence and `Cref` at the beginning.

1.3 Citations

Mathematical typesetting follows the principles established by [1], while modern computational approaches build on [2].

1.4 Tables and Figures

1.4.1 Professional Tables with Booktabs

Here's a professional table using the booktabs package:

Table 1.1: Comparison of Machine Learning Models

Model	Accuracy (%)	Training Time (min)
Linear Regression	78.5	2.3
Random Forest	85.2	15.7
Neural Network	92.1	120.4

1.4.2 Colored Tables

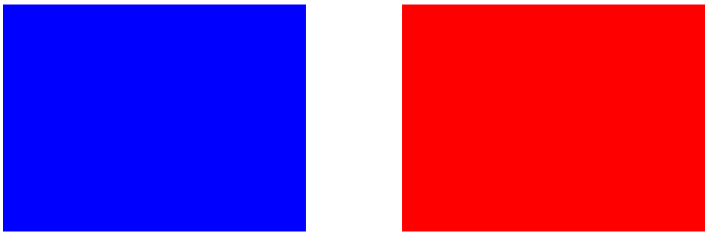
Using `xcolor` and `colortbl` packages for enhanced presentations:

1.4.3 Subfigures and Enhanced Captions

The subcaption package allows for complex figure arrangements:

Table 1.2: Performance Metrics with Color Coding

Method	Precision	Recall
Baseline	0.75	0.68
Our Method	**0.89**	**0.91**
State-of-art	0.82	0.85



(a) Training Loss (b) Validation Accuracy

Figure 1.1: Training Progress: a shows decreasing loss while b shows improving accuracy

1.4.4 Colors and Typography

The `xcolor` package enables rich text formatting:

- Important warnings can be highlighted
- Key concepts stand out
- Success indicators provide feedback
- Highlighted text draws attention

Mathematical expressions can also use colors:

$$\boldsymbol{W}\boldsymbol{x} + \boldsymbol{b} = \boldsymbol{y} \tag{1.26}$$

Note: The `parskip` package automatically improves paragraph spacing throughout the document.

Bibliography

[1] Donald E. Knuth. *The TeXbook*. Addison-Wesley, 1984.

[2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *Nature* 521.7553 (2015), pp. 436–444.