

Categorical Framework for Dimensional Analysis and System Classification

Abstract

This paper develops a categorical framework to formalize dimensional analysis using display categories, functors, and natural transformations. By introducing isometric and isomorphic invariances of dimensionless groups, we establish a method to classify physical systems into equivalence classes based on their behaviors. This framework unifies different physical domains and provides a systematic approach to analyze and predict system behaviors across various fields, including applications to machine learning models. We also explore the implications of linear dependence in dimension vector spaces on overfitting in modeling.

Contents

1	Introduction	2
2	Preliminaries	2
2.1	Basic Definitions in Category Theory	2
2.2	Dimensional Analysis and Dimensionless Groups	3
2.3	Dimension Vector Spaces	3
3	Display Categories and Dimensional Systems	3
3.1	Display Categories	3
3.2	Base Category of Dimensions	3
3.3	Variables and Morphisms	3
4	Functors and Natural Transformations in Dimensional Systems	3
4.1	Unit Systems as Functors	3
4.2	Natural Transformations as Unit Conversions	4
4.3	Invariance of Dimensionless Groups	4
5	Isometric and Isomorphic Invariances	4
5.1	Isometric Invariance	4
5.2	Isomorphic Invariance	4
6	Equivalence Classes via Dimensionless Groups	4
6.1	Classification of Systems	4
6.2	Formation of Equivalence Classes	4
7	Examples	5
7.1	Reynolds Number in Fluid Dynamics	5
7.2	Electrical-Mechanical System Analogy	5
8	Extension to Dynamical Systems	6
8.1	Differential Equations as Morphisms	6
8.2	Transducers as Isometric Morphisms	6

9	Overfitting and Linear Dependence in Dimension Vector Spaces	6
9.1	Linearly Dependent Units and Dimensional Redundancy	6
9.2	Example: Weight, Mass, and Acceleration	6
9.3	Implications for Modeling and Overfitting	7
9.4	Avoiding Dimensional Overfitting in Models	7
9.5	Analogies with Machine Learning	7
10	Applications to Machine Learning Models	8
10.1	Categorical Representation of Machine Learning Models	8
10.2	Dimensionless Groups in Machine Learning	8
10.3	Transfer Learning as a Functor	8
11	Conclusion	8

1 Introduction

Dimensional analysis is a fundamental tool in physics and engineering, providing insight into the relationships between physical quantities and allowing for the formation of dimensionless groups that characterize system behaviors. In this paper, we extend the traditional approach by applying category theory to dimensional analysis, introducing *display categories*, functors, and natural transformations to create a formal framework for classifying and comparing physical systems. We further explore applications of this framework in machine learning models and discuss how linear dependence in dimension vector spaces relates to overfitting in modeling.

2 Preliminaries

2.1 Basic Definitions in Category Theory

Definition 2.1 (Category Concepts Carnival). A *category* \mathcal{C} consists of:

- A class of *objects* $\text{Ob}(\mathcal{C})$.
- For every pair of objects $A, B \in \text{Ob}(\mathcal{C})$, a set of *morphisms* (or arrows) $\text{Hom}_{\mathcal{C}}(A, B)$.
- For every triple of objects $A, B, C \in \text{Ob}(\mathcal{C})$, a composition law $\circ : \text{Hom}_{\mathcal{C}}(B, C) \times \text{Hom}_{\mathcal{C}}(A, B) \rightarrow \text{Hom}_{\mathcal{C}}(A, C)$.
- For every object $A \in \text{Ob}(\mathcal{C})$, an identity morphism $\text{id}_A \in \text{Hom}_{\mathcal{C}}(A, A)$.

These must satisfy the associativity and identity axioms.

Definition 2.2 (Fantastic Functor). A *functor* $F : \mathcal{C} \rightarrow \mathcal{D}$ between categories \mathcal{C} and \mathcal{D} is a mapping that:

- Assigns to each object $A \in \text{Ob}(\mathcal{C})$ an object $F(A) \in \text{Ob}(\mathcal{D})$.
- Assigns to each morphism $f : A \rightarrow B$ in \mathcal{C} a morphism $F(f) : F(A) \rightarrow F(B)$ in \mathcal{D} .

Such that:

- $F(\text{id}_A) = \text{id}_{F(A)}$ for all $A \in \text{Ob}(\mathcal{C})$.
- $F(g \circ f) = F(g) \circ F(f)$ for all composable morphisms f, g in \mathcal{C} .

Definition 2.3 (Natural Transformation Treasure). Given two functors $F, G : \mathcal{C} \rightarrow \mathcal{D}$, a *natural transformation* $\eta : F \Rightarrow G$ assigns to each object $A \in \text{Ob}(\mathcal{C})$ a morphism $\eta_A : F(A) \rightarrow G(A)$ in \mathcal{D} , such that for every morphism $f : A \rightarrow B$ in \mathcal{C} , the following diagram commutes:

$$\begin{array}{ccc} F(A) & \xrightarrow{F(f)} & F(B) \\ \eta_A \downarrow & & \downarrow \eta_B \\ G(A) & \xrightarrow{G(f)} & G(B) \end{array}$$

2.2 Dimensional Analysis and Dimensionless Groups

Definition 2.4 (Dimension Delight). A *dimension* is a fundamental measure of a physical quantity, represented symbolically (e.g., mass $[M]$, length $[L]$, time $[T]$).

A *dimensioned quantity* is a physical variable associated with a combination of dimensions, expressed as $[M]^a[L]^b[T]^c$, where $a, b, c \in \mathbb{R}$.

Definition 2.5 (Dimensionless Dazzler). A *dimensionless group* is a combination of physical variables and constants that results in a quantity without units, i.e., all dimensions cancel out.

2.3 Dimension Vector Spaces

Definition 2.6 (Vector Space Voyage). The *dimension vector space* \mathcal{V}_D is a real vector space formed by the exponents of the fundamental dimensions. Each dimensioned quantity corresponds to a vector in \mathcal{V}_D .

3 Display Categories and Dimensional Systems

3.1 Display Categories

Definition 3.1 (Display Dynamo). A *display category* \mathcal{V} over a base category \mathcal{D} is a category where:

- Objects of \mathcal{V} are variables associated with dimensions in \mathcal{D} .
- Morphisms in \mathcal{V} represent relationships (e.g., equations, transformations) between these variables.
- There is a functor $p : \mathcal{V} \rightarrow \mathcal{D}$ called the *display functor*, which maps each variable to its associated dimension.

3.2 Base Category of Dimensions

Let \mathcal{D} be the base category where:

- $\text{Ob}(\mathcal{D})$ consists of fundamental dimensions $[M], [L], [T], \dots$
- Morphisms in \mathcal{D} represent dimension transformations (e.g., scaling, combination).

3.3 Variables and Morphisms

In the display category \mathcal{V} :

- Objects are physical variables (e.g., mass m , velocity v , force F) with associated dimensions via the functor p .
- Morphisms represent physical laws or relationships between variables (e.g., $F = ma$, $v = \frac{dx}{dt}$).

4 Functors and Natural Transformations in Dimensional Systems

4.1 Unit Systems as Functors

Let \mathcal{U}_{SI} and $\mathcal{U}_{\text{Imperial}}$ be categories representing the SI and Imperial unit systems, respectively.

Define functors:

$$\begin{aligned} F_{\text{SI}} : \mathcal{D} &\rightarrow \mathcal{U}_{\text{SI}}, \\ F_{\text{Imperial}} : \mathcal{D} &\rightarrow \mathcal{U}_{\text{Imperial}}. \end{aligned}$$

These functors map dimensions to units in each system.

4.2 Natural Transformations as Unit Conversions

Definition 4.1 (Natural Conversions). A natural transformation $\eta : F_{\text{SI}} \Rightarrow F_{\text{Imperial}}$ represents unit conversions between the SI and Imperial systems.

For each dimension $[D] \in \text{Ob}(\mathcal{D})$, $\eta_{[D]} : F_{\text{SI}}([D]) \rightarrow F_{\text{Imperial}}([D])$ is the conversion factor between units.

4.3 Invariance of Dimensionless Groups

Theorem 4.1 (Dimensionless Invariance Theorem). *Let Π be a dimensionless group formed from variables in \mathcal{V} associated with dimensions in \mathcal{D} . Then, under the natural transformation $\eta : F_{\text{SI}} \Rightarrow F_{\text{Imperial}}$, Π remains invariant.*

Proof. Since Π is dimensionless, it is constructed such that all dimensional units cancel out. Under the natural transformation η , each variable transforms according to unit conversion factors, but the ratios and products in Π ensure that these factors cancel. Therefore, Π is invariant under η . \square

5 Isometric and Isomorphic Invariances

5.1 Isometric Invariance

Definition 5.1 (Isometric Invariance). An *isometric transformation* in this context is a mapping that preserves quantitative relationships (e.g., ratios, magnitudes) between variables.

Two systems are *isometrically invariant* if there exists a transformation between them that preserves these quantitative relationships.

5.2 Isomorphic Invariance

Definition 5.2 (Isomorphic Invariance). An *isomorphism* between two categories \mathcal{C} and \mathcal{D} is a functor $F : \mathcal{C} \rightarrow \mathcal{D}$ that is invertible, i.e., there exists a functor $G : \mathcal{D} \rightarrow \mathcal{C}$ such that $G \circ F = \text{id}_{\mathcal{C}}$ and $F \circ G = \text{id}_{\mathcal{D}}$.

Two systems are *isomorphically invariant* if their categories are isomorphic, preserving the structural relationships between variables and equations.

6 Equivalence Classes via Dimensionless Groups

6.1 Classification of Systems

Definition 6.1 (Equivalent Invariants). An *equivalence class* of systems is a set of systems that share the same dimensionless groups and exhibit similar behaviors under corresponding conditions.

We denote an equivalence class \mathcal{E}_{Π} for systems sharing a dimensionless group Π .

6.2 Formation of Equivalence Classes

Theorem 6.1 (Equivalence Enchantment Theorem). *Systems that are isometrically or isomorphically invariant under transformations preserving their dimensionless groups belong to the same equivalence class \mathcal{E}_{Π} .*

Proof. Let systems S_1 and S_2 have categories \mathcal{C}_1 and \mathcal{C}_2 , respectively, with associated dimensionless groups Π_1 and Π_2 .

Suppose there exists:

- An isometric transformation ϕ between S_1 and S_2 preserving quantitative relationships.
- An isomorphism $F : \mathcal{C}_1 \rightarrow \mathcal{C}_2$ preserving structural relationships.

Since ϕ and F preserve the dimensionless groups (i.e., $\Pi_1 = \Pi_2$), and the systems exhibit similar behaviors under these transformations, they belong to the same equivalence class \mathcal{E}_{Π} . \square

7 Examples

7.1 Reynolds Number in Fluid Dynamics

Definition 7.1 (Reynolds Revelation). The *Reynolds number* Re is a dimensionless group defined as:

$$\text{Re} = \frac{\rho v L}{\mu},$$

where:

- ρ is the fluid density ($[M][L]^{-3}$).
- v is the fluid velocity ($[L][T]^{-1}$).
- L is a characteristic length ($[L]$).
- μ is the dynamic viscosity ($[M][L]^{-1}[T]^{-1}$).

Theorem 7.1 (Flow Familiarity Theorem). *Fluid systems with the same Reynolds number Re exhibit similar flow characteristics and belong to the same equivalence class \mathcal{E}_{Re} .*

Proof. The Reynolds number Re encapsulates the ratio of inertial forces to viscous forces in a fluid flow. Systems with the same Re have proportional relationships between these forces, leading to similar flow patterns (e.g., laminar or turbulent flow).

Through an isometric transformation preserving the Reynolds number, variables in one system can be scaled to those in another while maintaining the quantitative relationships. Therefore, these systems exhibit similar behaviors and belong to the same equivalence class \mathcal{E}_{Re} . \square

7.2 Electrical-Mechanical System Analogy

Definition 7.2 (Analogous Alliance). Consider the mechanical system governed by:

$$F(t) = m\ddot{x}(t) + c\dot{x}(t) + kx(t),$$

and the electrical circuit governed by:

$$V(t) = L\ddot{Q}(t) + R\dot{Q}(t) + \frac{1}{C}Q(t).$$

Theorem 7.2 (System Symmetry Theorem). *The mechanical and electrical systems are isomorphically invariant under a functorial mapping that preserves the structural form of their governing equations.*

Proof. Define a functor $F : \mathcal{C}_{\text{mech}} \rightarrow \mathcal{C}_{\text{elec}}$ such that:

$$\begin{aligned} F(x(t)) &= Q(t), \\ F(\dot{x}(t)) &= \dot{Q}(t), \\ F(\ddot{x}(t)) &= \ddot{Q}(t), \\ F(F(t)) &= V(t), \\ F(m) &= L, \\ F(c) &= R, \\ F(k) &= \frac{1}{C}. \end{aligned}$$

The morphisms (differential relationships) are preserved under F :

$$F(m\ddot{x}(t) + c\dot{x}(t) + kx(t)) = L\ddot{Q}(t) + R\dot{Q}(t) + \frac{1}{C}Q(t).$$

Since F is invertible (with inverse mapping electrical quantities back to mechanical ones), and it preserves the structural relationships, the two systems are isomorphic. Therefore, they are isomorphically invariant and can be classified into the same equivalence class based on their dimensionless group analogs. \square

8 Extension to Dynamical Systems

8.1 Differential Equations as Morphisms

Definition 8.1 (Dynamic Dance). In dynamical systems, the evolution of variables over time can be represented as morphisms in the display category.

Let \mathcal{V} be a display category where:

- Objects are state variables $x(t)$, $v(t)$, $a(t)$.
- Morphisms are differential equations relating these variables, e.g., $v(t) = \frac{dx(t)}{dt}$.

8.2 Transducers as Isometric Morphisms

Theorem 8.1 (Transducer Transformation Theorem). *An ideal transducer can be represented as an isometric morphism in the display category, preserving power between domains.*

Proof. Consider an electrical-to-mechanical transducer:

$$\text{Electrical power } P_{\text{elec}} = V(t)I(t), \quad \text{Mechanical power } P_{\text{mech}} = F(t)v(t).$$

An ideal transducer satisfies $P_{\text{elec}} = P_{\text{mech}}$. The morphism $T : \mathcal{V}_{\text{elec}} \rightarrow \mathcal{V}_{\text{mech}}$ maps electrical variables to mechanical ones while preserving power:

$$\begin{aligned} T(V(t)) &= F(t), \\ T(I(t)) &= v(t). \end{aligned}$$

Since power is preserved, the quantitative relationships are maintained, making T an isometric morphism. \square

9 Overfitting and Linear Dependence in Dimension Vector Spaces

9.1 Linearly Dependent Units and Dimensional Redundancy

Definition 9.1 (Linear Dependence Delight). In the dimension vector space \mathcal{V}_D , a set of dimension vectors $\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_n\}$ is said to be *linearly dependent* if there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_n$, not all zero, such that:

$$\alpha_1 \mathbf{d}_1 + \alpha_2 \mathbf{d}_2 + \dots + \alpha_n \mathbf{d}_n = \mathbf{0}.$$

Proposition 9.1 (Dimension Dependency Proposition). *If a physical quantity's dimension vector is a linear combination of other dimension vectors, then the corresponding unit is dimensionally redundant in modeling.*

Proof. Let \mathbf{d}_w be the dimension vector of a physical quantity w , and suppose $\mathbf{d}_w = \mathbf{d}_m + \mathbf{d}_a$, where \mathbf{d}_m and \mathbf{d}_a are the dimension vectors of quantities m and a , respectively.

This implies that $w = m \cdot a$, and the dimensions of w are fully determined by m and a . Including w alongside m and a introduces redundancy, as w does not provide additional independent information. This redundancy can complicate models without adding value, analogous to multicollinearity in machine learning. \square

9.2 Example: Weight, Mass, and Acceleration

Definition 9.2 (Weighty Wonders). Consider the relationship:

$$w = mg,$$

where:

- w is the weight (force) of an object.
- m is the mass of the object.
- g is the acceleration due to gravity.

The dimension vectors are:

$$\begin{aligned}\mathbf{d}_w &= [M][L][T]^{-2} = (1, 1, -2), \\ \mathbf{d}_m &= [M] = (1, 0, 0), \\ \mathbf{d}_g &= [L][T]^{-2} = (0, 1, -2).\end{aligned}$$

We observe that:

$$\mathbf{d}_w = \mathbf{d}_m + \mathbf{d}_g.$$

9.3 Implications for Modeling and Overfitting

Theorem 9.1 (Overfitting Overload Theorem). *Including linearly dependent units in a model can lead to overfitting due to dimensional redundancy, analogous to multicollinearity in machine learning.*

Proof. In modeling physical systems, including variables whose dimension vectors are linearly dependent introduces redundancy. This increases the complexity of the model without providing additional independent information.

In machine learning, multicollinearity (linear dependence among features) can lead to overfitting, where the model captures noise instead of the underlying relationship.

By analogy, including dimensionally redundant variables can cause the model to overfit to specific forms of the data, reducing its generalizability. Recognizing and removing or appropriately handling these redundancies can improve model performance and interpretability. \square

9.4 Avoiding Dimensional Overfitting in Models

Proposition 9.2 (Simplification Strategy Proposition). *To prevent overfitting due to dimensional redundancy, models should be simplified by:*

- *Identifying and removing linearly dependent units.*
- *Using fundamental variables to represent the system.*
- *Employing dimensionless groups to reduce the number of variables.*

Proof. By focusing on a set of dimensionally independent variables, we reduce the complexity of the model. Dimensionless groups combine variables in a way that captures the essential behavior of the system without redundancy. This approach aligns with principles of dimensional analysis and improves model generalization. \square

9.5 Analogies with Machine Learning

Definition 9.3 (Feature Finesse). In machine learning, *multicollinearity* refers to the presence of highly correlated (linearly dependent) features in the dataset.

Corollary 9.1.1 (Modeling Multicollinearity Corollary). *The presence of linearly dependent dimension vectors in physical modeling is analogous to multicollinearity in machine learning models.*

Proof. In both cases, linear dependence among variables (features or dimension vectors) introduces redundancy. This can lead to overfitting, where the model becomes overly complex and sensitive to fluctuations in the data. By addressing linear dependence, we simplify the model and enhance its predictive capabilities. \square

10 Applications to Machine Learning Models

10.1 Categorical Representation of Machine Learning Models

Definition 10.1 (Model Morphism Magic). A *machine learning model* can be represented as a category \mathcal{M} where:

- Objects are layers or components of the model (e.g., neurons, layers in a neural network).
- Morphisms are the functions or mappings between layers (e.g., activation functions, weight matrices).

10.2 Dimensionless Groups in Machine Learning

Definition 10.2 (Learning Landscape Luminary). A *dimensionless group* in machine learning could be a hyperparameter ratio that influences model behavior, such as the ratio of learning rate to batch size.

Theorem 10.1 (Hyperparameter Harmony Theorem). *Models sharing the same dimensionless hyperparameter groups exhibit similar training dynamics and can be classified into the same equivalence class.*

Proof. Dimensionless hyperparameter groups encapsulate the relative influence of different training parameters on model behavior. By preserving these ratios, models can be scaled while maintaining similar training dynamics.

Therefore, models with the same dimensionless hyperparameter groups can be transformed into one another via isometric transformations, and exhibit similar behaviors, belonging to the same equivalence class. \square

10.3 Transfer Learning as a Functor

Definition 10.3 (Transfer Transformation). *Transfer learning* can be viewed as a functor $T : \mathcal{M}_1 \rightarrow \mathcal{M}_2$, mapping a pre-trained model \mathcal{M}_1 to a new model \mathcal{M}_2 for a different but related task.

Theorem 10.2 (Transfer Learning Functoriality Theorem). *If the transfer learning functor T preserves the structural and quantitative relationships (i.e., isomorphic and isometric invariances), then the performance on the new task can be predicted based on the original model's performance.*

Proof. By viewing transfer learning as a functor that maps models while preserving key relationships, we can ensure that the essential features learned in \mathcal{M}_1 are retained in \mathcal{M}_2 .

If T preserves invariances, then \mathcal{M}_2 benefits from the learned representations in \mathcal{M}_1 , leading to similar performance characteristics. Thus, understanding T allows us to predict and analyze \mathcal{M}_2 based on \mathcal{M}_1 . \square

11 Conclusion

We have developed a categorical framework that formalizes dimensional analysis using display categories, functors, and natural transformations. By introducing isometric and isomorphic invariances of dimensionless groups, we can classify physical systems into equivalence classes based on their behaviors. We have explored how linear dependence in dimension vector spaces can lead to overfitting due to dimensional redundancy, drawing parallels with multicollinearity in machine learning. This approach unifies different physical domains and enhances our ability to analyze and predict system behaviors across various fields, including machine learning models. The framework provides a foundation for further exploration of categorical structures in complex systems.