Семантический Анализ: Математический Формализм для Операций над Нейровесовыми Полями

Белоусов Р.С., *ORCID: 0009-0009-7262-633X*
Независимый исследователь, Российская Федерация, г. Москва.

Аннотация. Данная работа вводит новую ветвь математики — **Семантический Анализ** (СА) — формальный аппарат для операций над объектами, моделирующими смысл, а не синтаксис данных. Исходными объектами СА являются **семантические объекты** $S = (z, \Sigma)$, пары, состоящие из семантического ядра z (низкоразмерного вектора-признака) и ковариационной матрицы Σ, количественно определяющей «уверенность» или «семантический разброс» модели. Мы постулируем систему аксиом, определяющих свойства семантического пространства S, и вводим новый класс операций: **семантическая суперпозиция ($\oplus$)**, **семантическая интерполяция ($\circledast$)** и **семантический градиент (∇_s)**. Эти операции, в отличие от операций линейной алгебры, некоммутативны, неассоциативны и учитывают байесовскую природу лежащих в их основе моделей. Доказываются ключевые теоремы, связывающие СА с теорией информации (через расхождение Кульбака-Лейблера) и теорией поля. СА предоставляет строгий математический фундамент для теорий, подобных теории Нейровесовых Полей (NWF), открывая путь к созданию систем ИИ, оперирующих не данными, а смыслами.

Ключевые слова: семантический анализ; нейровесовые поля; байесовский вывод; теория информации; семантические операции; некоммутативная алгебра.

Для цитирования: Белоусов Р.С. Семантический Анализ: Математический Формализм для Операций над Нейровесовыми Полями.

---

1. Введение: Мотивация и Постановка Задачи

Современные парадигмы машинного обучения (ML) и информационного поиска оперируют в основном синтаксическими представлениями данных: битовыми строками, тензорами, векторами в R^n. Операции над ними (сложение, умножение, вычисление косинусного расстояния) лишены семантического контекста. Семантика привносится внешними моделями post factum.

Теория Нейровесовых Полей (NWF) [1] предлагает альтернативу: данные хранятся как параметры θ нейронной сети, полученные в результате байесовского вывода. Это представление инкапсулирует семантику. Однако NWF требуют нового математического языка для описания операций над этими параметризованными моделями.

Данная работа заполняет эту нишу, вводя **Семантический Анализ** — формальную систему для операций над **семантическими объектами** $S_i = (z_i, \Sigma_i)$, где:

- $z_i \in R^d$ — семантическое ядро (латентный вектор).
- $\Sigma_i \in R^{(d \times d)}$ — ковариационная матрица, $\Sigma_i \succ 0$ (положительно определённая).

Матрица Σ_i формализует понятие «уверенности модели». Малая Σ (высокая точность) означает уверенность в семантике z; большая Σ (низкая точность) — размытое, неопределённое значение.

2. Аксиоматика Семантического Пространства

Пусть S — множество всех семантических объектов. Определим на нём структуру **Семантического Пространства**.

Аксиома 1 (Аксиома Данных-как-Модели). Для любого набора данных D существует непустое множество семантических объектов $S_D \subset S$, являющихся его байесовскими моделями в рамках некоторой гипотезы H.

Аксиома 2 (Аксиома Метрики Уверенности). На пространстве S задана функция расстояния $d_S: S \times S \to R^+ \cup \{0\}$, являющаяся модификацией расхождения Кульбака-Лейблера (KL-divergence) между двумя гауссовскими распределениями $N(z_i, \Sigma_i)$ и $N(z_j, \Sigma_j)$:

$$d_S(S_i, S_j) = D_{KL}(N_i \| N_j) = 1/2 \left[tr(\Sigma_j^{-1} \Sigma_i) + (z_j - z_i)^T \Sigma_j^{-1} (z_j - z_i) - d + \ln(|\Sigma_j| / |\Sigma_i|) \right] \quad (1)$$

Эта метрика измеряет «семантическую близость» с учётом уверенности моделей.

Аксиома 3 (Аксиома Нулевого Объекта). Существует **нулевой семантический объект** $S_0 = (0, I)$, где I — единичная матрица. Этот объект представляет полное отсутствие информации или максимальную энтропию.

3. Базовые Операции Семантического Анализа

3.1. Семантическая Суперпозиция (Semantic Superposition) ⊕

Определение 1. Операция суперпозиции $\oplus: S \times S \to S$ двух объектов $S_1 = (z_1, \Sigma_1)$ и $S_2 = (z_2, \Sigma_2)$ определяется как байесовское слияние их информационных содержаний:

$$S_{total} = S_1 \oplus S_2 = (z_{total}, \Sigma_{total})$$

$$z_{total} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} (\Sigma_1^{-1} z_1 + \Sigma_2^{-1} z_2) \quad (2)$$
$$\Sigma_{total} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} \quad (3)$$

Свойства:

1. **Некоммутативность:** $S_1 \oplus S_2 \neq S_2 \oplus S_1$ (порядок информации важен).
2. **Неассоциативность:** $(S_1 \oplus S_2) \oplus S_3 \neq S_1 \oplus (S_2 \oplus S_3)$.
3. **Нейтральный элемент:** $S \oplus S_0 = S_0 \oplus S = S$.
4. **Идемпотентность (приближённая):** $S \oplus S \approx S$.

Интерпретация: Суперпозиция объединяет знания из двух источников, взвешивая их по их уверенности (Σ^{-1} — информационная матрица). Более уверенный источник вносит больший вклад в итог.

3.2. Семантическая Интерполяция (Semantic Interpolation) ⊛

Определение 2. Операция интерполяции $\circledast: R \times S \to S$ и $\oplus\_\circledast: S \times S \to S$ позволяет находить промежуточные смыслы. Для скаляра $\alpha \in [0, 1]$ и объекта $S = (z, \Sigma)$:
$\alpha \circledast S = (\alpha z, \alpha \Sigma + (1-\alpha) I)$ (4)

Для двух объектов интерполяция определяется через их суперпозицию с весами:
$S_{new} = (\alpha) \circledast S_1 \oplus ((1-\alpha)) \circledast S_2$ (5)

Интерпретация: Операция создаёт плавный переход между двумя концепциями. Уверенность итогового объекта максимальна при $\alpha=0.5$ (наибольшая неопределённость точки перехода) и минимальна на исходных объектах.

4. Дифференциальные Операторы в Семантическом Пространстве

Для описания динамики семантических полей введём дифференциальные операторы.

4.1. Семантический Градиент (Semantic Gradient) ∇_s

Определение 3. Пусть $U: S \to R$ — семантический потенциал (скалярная функция на S). **Семантический градиент** $\nabla_s U$ — это оператор, возвращающий вектор в R^d, указывающий направление наискорейшего увеличения потенциала U в пространстве ядер z:

$\nabla_s U(S) = [\partial U/\partial z_1, \partial U/\partial z_2, ..., \partial U/\partial z_d]^T$ (6)

Теорема 1 (О Семантической Силе). Сила семантического взаимодействия $F_{i\to j}$, с которой объект S_i действует на объект S_j, равна семантическому градиенту от интеграла перекрытия их потенциалов $S(i,j)$ (см. [1], Теорема 3.2), взятого с обратным знаком:

$F_{i\to j} = -\nabla_{z_j} U = k * \nabla_{z_j} S(i,j)$, где $S(i,j) = \int \varphi_i(r) \varphi_j(r) dr$ (7)

Для гауссовых ядер $\varphi_i(r) \sim N(z_i, \Sigma_i)$ сила может быть вычислена аналитически.

4.2. Семантический Лапласиан (Semantic Laplacian) Δ_s

Определение 4. Семантический лапласиан — оператор второго порядка, измеряющий расходимость семантического поля:
$\Delta_s U(S) = tr(\Sigma * H(U))$ (8)
где $H(U)$ — гессиан функции U по z, а Σ — ковариация объекта, в точке которого вычисляется оператор. Этот оператор учитывает "разброс" семантики объекта.

5. Приложения: Семантический Поиск и Динамическое Обучение

5.1. Семантический Поиск

Запрос Q к системе NWF кодируется в семантический объект $S_q = (z_q, \Sigma_q)$. Система находит в базе объект S_d, минимизирующий расстояние $d_S(S_q, S_d)$ (Ур. 1). Это эффективнее косинусной меры, так как запрос "кошка" будет ближе к "животное", если Σ_q велика (размытый запрос), и ближе к "мейн-кун", если Σ_q мала (конкретный запрос).

5.2. Динамическое Обучение как Суперпозиция

Процесс обучения на новом батче данных `D_new` формализуется как итеративная суперпозиция:
$S_{(t+1)} = S_t \oplus S_{new}$ (9)
где `S_new` — объект, полученный кодированием `D_new`. Это реализует онлайн-байесовский вывод. Данная формализация позволяет строго анализировать сходимость обучения и "катастрофическое забывание".

6. Заключение и Направления Развития

Введённый формализм **Семантического Анализа** предоставляет мощный и строгий математический язык для описания операций над смыслами. Его ключевые особенности — учёт уверенности модели через Σ и некоммутативность операций — отражают фундаментальные свойства семантической информации.

Направления будущих исследований:

1. **Теорема о полноте:** Доказательство, что система операторов $\{\oplus, \circledast, \nabla_s, \Delta_s\}$ является полной для описания любых семантических преобразований.
2. **Топология семантического пространства:** Исследование сходимости последовательностей семантических объектов и компактности подмножеств S.
3. **Семантическое интегрирование:** Определение интеграла от функции по семантическому полю $\int f(S)\, dS$.
4. **Связь с некоммутативной геометрией:** Исследование глубоких связей между алгеброй операторов СА и некоммутативной геометрией Алена Конна.

Семантический Анализ закладывает основу для математики будущих семантически-ориентированных вычислительных систем.

---

1. Белоусов Р.С. Нейровесовые поля: Теория Семантического Континуума для Хранения и Обработки Информации. 2023.

2. Кульбак С., Лейблер Р.А. Об информации и достаточности // Анналы математической статистики. 1951. Т. 22. № 1. С. 79-86. (Kullback, S., Leibler, R.A. On Information and Sufficiency. The Annals of Mathematical Statistics, 1951, vol. 22, no. 1, pp. 79-86.)

3. Петерсен К.Б., Педерсен М.С. Поваренная книга по матрицам. Технический университет Дании, 2012. (Petersen, K. B., & Pedersen, M. S. The Matrix Cookbook. Technical University of Denmark, 2012.)

4. Сарккя С. Байесовская фильтрация и сглаживание. Издательство Кембриджского университета, 2013. Т. 3. (Sarkka, S. Bayesian Filtering and Smoothing. Cambridge university press, 2013, vol. 3.)

5. Конн А. Некоммутативная геометрия. Academic Press, 1994. (Connes, A. Noncommutative Geometry. Academic Press, 1994.)

6. Фристон К. Принцип свободной энергии: унифицированная теория мозга? // Nature Reviews Neuroscience. 2010. Т. 11. № 2. С. 127-138. (Friston, K. The free-energy principle: a unified brain theory?. Nature Reviews Neuroscience, 2010, vol. 11, no. 2, pp. 127-138.)

7. Гудфеллоу А., Бенджио И., Курвиль А. Глубокое обучение. MIT press, 2016. (Goodfellow, I., Bengio, Y., & Courville, A. Deep learning. MIT press, 2016.)

8. Джейнс Э.Т. Теория вероятностей: логика науки. Издательство Кембриджского университета, 2003. (Jaynes, E.T. Probability Theory: The Logic of Science. Cambridge university press, 2003.)

9. Малков Ю.А., Яшунин Д.А. Эффективный и надежный приближенный поиск ближайшего соседа с использованием графов Иерархичных Навигируемых Малых Миров (HNSW) // IEEE Transactions по анализу образов и машинному интеллекту. 2018. Т. 42. № 4. С. 824-836. (Malkov, Y. A., & Yashunin, D. A. Efficient and robust approximate nearest neighbor search using Hierarchical Navigable Small World graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, vol. 42, no. 4, pp. 824-836.)

10. Милденхолл Б., Шринивасан П.П., Танчик М. и др. NeRF: Представление сцен в виде нейронных полей излучения для синтеза видов // Communications of the ACM. 2021. Т. 65. № 1. С. 99-106. (Mildenhall, B., Srinivasan, P. P., Tancik, M., et al. NeRF: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 2021, vol. 65, no. 1, pp. 99-106.)

Semantic Calculus: A Mathematical Formalism for Operations on Neural Weight Fields

Belousov R.S., *ORCID: 0009-0009-7262-633X*
Independent researcher, Moscow, Russian Federation.

Abstract. This work introduces a new branch of mathematics—**Semantic Calculus** (SC)—a formal apparatus for operations on objects that model meaning, not data syntax. The fundamental objects of SC are **semantic objects** $S = (z, \Sigma)$, pairs consisting of a semantic kernel z (a low-dimensional feature vector) and a covariance matrix Σ, which quantitatively defines the "confidence" or "semantic spread" of the model. We postulate a system of axioms defining the properties of the semantic space S and introduce a new class of operations: **semantic superposition ($\oplus$)**, **semantic interpolation ($\circledast$)**, and **semantic gradient (∇_s)**. Unlike linear algebra operations, these are non-commutative, non-associative, and account for the Bayesian nature of the underlying models. Key theorems linking SC to information theory (via Kullback-Leibler divergence) and field theory are proven. SC provides a rigorous mathematical foundation for theories like Neural Weight Fields (NWF), paving the way for building AI systems that operate not on data, but on meanings.

Keywords: semantic calculus; neural weight fields; bayesian inference; information theory; semantic operations; non-commutative algebra.

For citation: Belousov R.S. Semantic Calculus: A Mathematical Formalism for Operations on Neural Weight Fields.

1. Introduction: Motivation and Problem Statement

Modern paradigms of machine learning (ML) and information retrieval primarily operate on syntactic representations of data: bitstrings, tensors, vectors in R^n. Operations on them (addition, multiplication, cosine similarity calculation) lack semantic context. Semantics is introduced post factum by external models.

The theory of Neural Weight Fields (NWF) [1] offers an alternative: data is stored as parameters θ of a neural network obtained through Bayesian inference. This representation encapsulates semantics. However, NWF requires a new mathematical language to describe operations on these parameterized models.

This work fills this niche by introducing **Semantic Calculus**—a formal system for operations on **semantic objects** $S_i = (z_i, \Sigma_i)$, where:

- $z_i \in R^d$ is the semantic kernel (latent vector).
- $\Sigma_i \in R^{(d \times d)}$ is the covariance matrix, $\Sigma_i \succ 0$ (positive definite).

The matrix Σ_i formalizes the concept of "model confidence." A small Σ (high precision) means confidence in the semantics of z; a large Σ (low precision) means a vague, uncertain meaning.

2. Axiomatics of the Semantic Space

Let S be the set of all semantic objects. We define a **Semantic Space** structure on it.

Axiom 1 (Data-as-Model Axiom). For any dataset D, there exists a non-empty set of semantic objects $S_D \subset S$ that are its Bayesian models under some hypothesis H.

Axiom 2 (Confidence Metric Axiom). A distance function $d_S: S \times S \to R^+ \cup \{0\}$ is defined on the space S, which is a modification of the Kullback-Leibler divergence (KL-divergence) between two Gaussian distributions $N(z_i, \Sigma_i)$ and $N(z_j, \Sigma_j)$:

$$d_S(S_i, S_j) = D_{KL}(N_i \| N_j) = 1/2 \left[tr(\Sigma_j^{-1} \Sigma_i) + (z_j - z_i)^T \Sigma_j^{-1} (z_j - z_i) - d + \ln(|\Sigma_j| / |\Sigma_i|) \right] \quad (1)$$

This metric measures "semantic proximity" taking into account the confidence of the models.

Axiom 3 (Zero Object Axiom). There exists a **zero semantic object** $S_0 = (0, I)$, where I is the identity matrix. This object represents a complete lack of information or maximum entropy.

3. Basic Operations of Semantic Calculus

3.1. Semantic Superposition ⊕

Definition 1. The superposition operation $\oplus: S \times S \to S$ of two objects $S_1 = (z_1, \Sigma_1)$ and $S_2 = (z_2, \Sigma_2)$ is defined as the Bayesian fusion of their informational content:

$$S_{total} = S_1 \oplus S_2 = (z_{total}, \Sigma_{total})$$

$$z_{total} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} (\Sigma_1^{-1} z_1 + \Sigma_2^{-1} z_2) \quad (2)$$
$$\Sigma_{total} = (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} \quad (3)$$

Properties:

1. **Non-commutativity:** $S_1 \oplus S_2 \neq S_2 \oplus S_1$ (the order of information matters).
2. **Non-associativity:** $(S_1 \oplus S_2) \oplus S_3 \neq S_1 \oplus (S_2 \oplus S_3)$.
3. **Neutral element:** $S \oplus S_0 = S_0 \oplus S = S$.
4. **Idempotency (approximate):** $S \oplus S \approx S$.

Interpretation: Superposition combines knowledge from two sources, weighting them by their confidence (Σ^{-1} is the precision matrix). A more confident source contributes more to the result.

3.2. Semantic Interpolation ⊛

Definition 2. The interpolation operation $\circledast: R \times S \to S$ and $\oplus_\circledast: S \times S \to S$ allows finding intermediate meanings. For a scalar $\alpha \in [0, 1]$ and an object $S = (z, \Sigma)$:
$$\alpha \circledast S = (\alpha z, \alpha \Sigma + (1-\alpha) I) \quad (4)$$

For two objects, interpolation is defined via their superposition with weights:
$$S_{new} = (\alpha) \circledast S_1 \oplus ((1-\alpha)) \circledast S_2 \quad (5)$$

Interpretation: The operation creates a smooth transition between two concepts. The confidence of the resulting object is highest at $\alpha=0.5$ (greatest uncertainty of the transition point) and lowest at the original objects.

4. Differential Operators in Semantic Space

To describe the dynamics of semantic fields, we introduce differential operators.

4.1. Semantic Gradient ∇_s

Definition 3. Let $U: S \to R$ be a semantic potential (a scalar function on S). The **semantic gradient** $\nabla_s U$ is an operator returning a vector in R^d, indicating the direction of the steepest increase of the potential U in the kernel space z:

$$\nabla_s U(S) = [\ \partial U/\partial z_1, \partial U/\partial z_2, \ldots, \partial U/\partial z_d\ ]^T \quad (6)$$

Theorem 1 (On Semantic Force). The force of semantic interaction $F_{i\to j}$ with which object S_i acts on object S_j is equal to the semantic gradient of the overlap integral of their potentials $S(i,j)$ (see [1], Theorem 3.2), taken with a negative sign:

$$F_{i\to j} = -\nabla_{z_j} U = k * \nabla_{z_j} S(i,j), \text{ where } S(i,j) = \int \varphi_i(r)\, \varphi_j(r)\, dr \quad (7)$$

For Gaussian kernels $\varphi_i(r) \sim N(z_i, \Sigma_i)$, the force can be computed analytically.

4.2. Semantic Laplacian Δ_s

Definition 4. The **semantic Laplacian** is a second-order operator measuring the divergence of a semantic field:

$\Delta_s U(S) = tr(\Sigma * H(U))$ (8)
where $H(U)$ is the Hessian of the function U with respect to z, and Σ is the covariance of the object at which the operator is computed. This operator accounts for the "spread" of the object's semantics.

5. Applications: Semantic Search and Dynamic Learning

5.1. Semantic Search

A query Q to an NWF system is encoded into a semantic object $S_q = (z_q, \Sigma_q)$. The system finds the object S_d in the database that minimizes the distance $d_S(S_q, S_d)$ (Eq. 1). This is more effective than cosine similarity because the query "cat" will be closer to "animal" if Σ_q is large (vague query), and closer to "Maine Coon" if Σ_q is small (specific query).

5.2. Dynamic Learning as Superposition

The process of learning on a new batch of data D_{new} is formalized as an iterative superposition:
$S_{(t+1)} = S_t \oplus S_{new}$ (9)
where S_{new} is the object obtained by encoding D_{new}. This implements online Bayesian inference. This formalization allows for rigorous analysis of learning convergence and "catastrophic forgetting."

6. Conclusion and Future Research Directions

The introduced formalism of **Semantic Calculus** provides a powerful and rigorous mathematical language for describing operations on meanings. Its key features—accounting for model confidence via Σ and the non-commutativity of operations—reflect the fundamental properties of semantic information.

Future research directions:

1. **Completeness theorem:** Proving that the system of operators $\{\oplus, \circledast, \nabla_s, \Delta_s\}$ is complete for describing any semantic transformations.
2. **Topology of semantic space:** Investigating the convergence of sequences of semantic objects and the compactness of subsets of S.
3. **Semantic integration:** Defining the integral of a function over a semantic field $\int f(S)\, dS$.
4. **Connection to non-commutative geometry:** Exploring deep connections between the algebra of SC operators and Alain Connes' non-commutative geometry.

Semantic Calculus lays the foundation for the mathematics of future semantically-oriented computing systems.

References

1. Belousov, R.S. Neural Weight Fields: Theory of Semantic Continuum for Information Storage and Processing. 2023.

2. Kullback, S., Leibler, R.A. On Information and Sufficiency. The Annals of Mathematical Statistics, 1951, vol. 22, no. 1, pp. 79-86.

3. Petersen, K. B., & Pedersen, M. S. The Matrix Cookbook. Technical University of Denmark, 2012.

4. Sarkka, S. Bayesian Filtering and Smoothing. Cambridge university press, 2013, vol. 3.

5. Connes, A. Noncommutative Geometry. Academic Press, 1994.

6. Friston, K. The free-energy principle: a unified brain theory?. Nature Reviews Neuroscience, 2010, vol. 11, no. 2, pp. 127-138.

7. Goodfellow, I., Bengio, Y., & Courville, A. Deep Learning. MIT press, 2016.

8. Jaynes, E.T. Probability Theory: The Logic of Science. Cambridge university press, 2003.

9. Malkov, Y. A., & Yashunin, D. A. Efficient and robust approximate nearest neighbor search using Hierarchical Navigable Small World graphs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, vol. 42, no. 4, pp. 824-836.

10. Mildenhall, B., Srinivasan, P. P., Tancik, M., et al. NeRF: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 2021, vol. 65, no. 1, pp. 99-106.