GRADIENT FORMULA CALCULUS

Gradient formula calculus is a fundamental concept in the field of mathematics, particularly in multivariable calculus. Understanding the gradient is essential for various applications, including optimization problems, physics, and computer graphics. This article will explore the gradient formula in detail, its mathematical derivation, applications, and significance in different fields. We will provide a comprehensive overview of how the gradient formula is utilized in calculus, making it easier for students and professionals alike to grasp its importance. By the end of this article, readers will have a thorough understanding of the gradient formula, its computation, and its practical applications.

- Introduction to the Gradient Formula
- Mathematical Definition of the Gradient
- Calculating the Gradient
- Applications of the Gradient in Various Fields
- Gradient in Optimization Problems
- Conclusion

Introduction to the Gradient Formula

The gradient formula is a vector that represents the direction and rate of the steepest ascent of a function. It is a crucial concept in calculus, particularly when dealing with functions of multiple variables. The gradient is denoted by the symbol ∇ (nabla) and is calculated from the partial derivatives of a function. The fundamental idea of the gradient is that it provides information about how a function changes as one moves through its domain.

In simpler terms, the gradient indicates how much a function increases or decreases in different directions. For instance, in a three-dimensional space, the gradient can show which direction to move to increase the height of a surface defined by a function. Understanding the gradient allows for better analysis of functions, particularly in optimization and various scientific applications.

Mathematical Definition of the Gradient

The gradient of a scalar function \(f(x, y, z) \) is defined mathematically as the vector of its partial derivatives. For a function of three variables, the gradient is represented as follows:

The gradient is defined as:

$$\nabla f = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$$

This formula indicates that the gradient consists of three components, each representing the rate of change of the function with respect to one of the input variables. The gradient is a vector that points in the direction of the steepest ascent of the function.

The Gradient in Two Dimensions

For functions of two variables, the gradient can be expressed as:

▯ f = (▯ f/▯ x, ▯ f/▯ y)

This two-dimensional representation is significant in many applications, including physics and engineering, where functions often depend on two variables such as \(x \) and \(y \).

Calculating the Gradient

To calculate the gradient of a function, follow these steps:

1. Identify the function \(f(x, y, z) \) you wish to analyze.
2. Compute the partial derivatives of the function with respect to each variable.
3. Form the gradient vector using the calculated partial derivatives.

Example Calculation

Let's consider the function \(f(x, y) = x^2 + y^2 \). We will compute the gradient.

1. Compute the partial derivatives:
- \(▯ f/▯ x = 2x \)
- \(▯ f/▯ y = 2y \)

2. Form the gradient vector:
- \(▯ f = (2x, 2y) \)

This gradient indicates that at any point \((x, y) \), the function \(f \) increases most rapidly in the direction of the vector \((2x, 2y) \).

Applications of the Gradient in Various Fields

The gradient has numerous applications across different domains. Here are a few notable examples:

- **Physics:** In physics, the gradient is used to describe various phenomena, such as the electric field, which is the gradient of the electric potential. It helps in understanding how forces act on charged particles.
- **Economics:** Economists use gradients to analyze cost functions and optimize production processes, ensuring maximum efficiency.
- **Machine Learning:** In machine learning, gradients are essential for optimization algorithms, particularly in training models through gradient descent methods.
- **Computer Graphics:** The gradient is utilized in rendering techniques, enhancing the realism of images by

simulating light behavior.

Gradient in Optimization Problems

Optimization problems often involve finding the maximum or minimum of a function. The gradient plays a crucial role in these scenarios:

In optimization, the gradient can indicate the direction to move to increase or decrease the function value. If the gradient at a point is zero, it indicates a potential local maximum, minimum, or saddle point. This is known as the critical point.

Gradient Descent Algorithm

One of the most common optimization techniques is the gradient descent algorithm, which iteratively moves towards the minimum of a function. The steps involved in gradient descent are:

1. Start with an initial guess x_0.
2. Compute the gradient $\nabla f(x_n)$ at the current point.
3. Update the current point using the formula:

   $x_{n+1} = x_n - a \nabla f(x_n)$

4. Repeat until convergence.

The learning rate a determines how large of a step to take in the direction of the gradient. Proper selection of a is critical for the success of the algorithm.

Conclusion

The gradient formula calculus is a pivotal concept in understanding how functions behave in multi-dimensional spaces. By representing the rate and direction of change, the gradient serves as a fundamental tool in various mathematical and applied fields. Through its calculation, one can glean critical insights about function behavior, optimize processes, and analyze real-world phenomena effectively. Mastery of the gradient and its applications not only enhances mathematical skills but also provides a robust framework for tackling complex scientific and engineering challenges.

Q: What is the gradient of a function?

A: The gradient of a function is a vector that consists of its partial derivatives. It indicates the direction and rate of the steepest ascent of the function.

Q: How do you compute the gradient?

A: To compute the gradient, identify the function, calculate the partial derivatives with respect to each variable, and form a vector from these derivatives.

Q: Why is the gradient important in optimization?

A: The gradient is crucial in optimization because it shows the direction to move to increase or decrease a function's value, helping to locate maximum or minimum points.

Q: What is gradient descent?

A: Gradient descent is an iterative optimization algorithm used to minimize a function by moving in the direction of the negative gradient.

Q: Can the gradient be used in machine learning?

A: Yes, the gradient is extensively used in machine learning for optimizing models, particularly in training algorithms that rely on gradient descent.

Q: What does it mean if the gradient is zero?

A: If the gradient is zero at a point, it indicates a critical point, which could be a local maximum, minimum, or saddle point of the function.

Q: How does the learning rate affect gradient descent?

A: The learning rate determines the size of the steps taken in the gradient direction. A small learning rate may lead to slow convergence, while a large learning rate can cause overshooting and divergence.

Q: In which fields is the gradient applied?

A: The gradient is applied in various fields, including physics, economics, machine learning, engineering, and computer graphics.

Q: What is the geometric interpretation of the gradient?

A: Geometrically, the gradient represents a vector pointing in the direction of the steepest ascent on the surface defined by the function, with its magnitude indicating the rate of increase.

Gradient Formula Calculus

Find other PDF articles:
https://ns2.kelisto.es/suggest-articles-01/files?trackid=VXA54-7301&title=how-to-write-a-qualitative-systematic-review.pdf

gradient formula calculus: Calculus Howard Anton, Irl C. Bivens, Stephen Davis, 2021-11-02 In the newly revised Twelfth Edition of Calculus, an expert team of mathematicians delivers a rigorous and intuitive exploration of calculus, introducing polynomials, rational functions, exponentials, logarithms, and trigonometric functions late in the text. Using the Rule of Four, the authors present mathematical concepts from verbal, algebraic, visual, and numerical points of view. The book includes numerous exercises, applications, and examples that help readers learn and retain the concepts discussed within.

gradient formula calculus: The Data Science Handbook Field Cady, 2025-01-09 Practical, accessible guide to becoming a data scientist, updated to include the latest advances in data science and related fields. Becoming a data scientist is hard. The job focuses on mathematical tools, but also demands fluency with software engineering, understanding of a business situation, and deep understanding of the data itself. This book provides a crash course in data science, combining all the necessary skills into a unified discipline. The focus of The Data Science Handbook is on practical applications and the ability to solve real problems, rather than theoretical formalisms that are rarely needed in practice. Among its key points are: Readers of the third edition of Construction Graphics will also find: An emphasis on software engineering and coding skills, which play a significant role in most real data science problems. Extensive sample code, detailed discussions of important libraries, and a solid grounding in core concepts from computer science (computer architecture, runtime complexity, programming paradigms, etc.) A broad overview of important mathematical tools, including classical techniques in statistics, stochastic modeling, regression, numerical optimization, and more. Extensive tips about the practical realities of working as a data scientist, including understanding related jobs functions, project life cycles, and the varying roles of data science in an organization. Exactly the right amount of theory. A solid conceptual foundation is required for fitting the right model to a business problem, understanding a tool's limitations, and reasoning about discoveries. Data science is a quickly evolving field, and the 2nd edition has been updated to reflect the latest developments, including the revolution in AI that has come from Large Language Models and the growth of ML Engineering as its own discipline. Much of data science has become a skillset that anybody can have, making this book not only for aspiring data scientists, but also for professionals in other fields who want to use analytics as a force multiplier in their organization.

gradient formula calculus: Mathematics Keith Devlin, 1996-12-15 To most people, mathematics means working with numbers. But as Keith Devlin shows in Mathematics: The Science of Patterns, this definition has been out of date for nearly 2,500 years. Mathematicians now see their work as the study of patterns—real or imagined, visual or mental, arising from the natural world or from within the human mind. Using this basic definition as his central theme, Devlin explores the patterns of counting, measuring, reasoning, motion, shape, position, and prediction, revealing the powerful influence mathematics has over our perception of reality. Interweaving historical highlights and current developments, and using a minimum of formulas, Devlin celebrates the precision, purity, and elegance of mathematics.

gradient formula calculus: *Business Analytics in Context* Gareth Woods, 2020-09-08 An essential resource for any student undertaking a business/management degree. An engaging and accessible text structured around the learning journey from understanding basic concepts, such as algebra, through to applying more advanced techniques including differentiation and optimisation.

gradient formula calculus: *Basic Physics: A Formula Handbook* N.B. Singh, Basic Physics: A Formula Handbook is an essential and user-friendly guide that distills fundamental principles of physics into a concise collection of formulas. Tailored for students, educators, and enthusiasts in the field, this handbook covers key equations spanning mechanics, electromagnetism, thermodynamics, and more. With clarity and simplicity, the book provides quick references for solving physics problems and understanding core concepts. Whether preparing for exams or seeking a quick review, this handbook serves as an invaluable resource, offering a straightforward and comprehensive approach to navigating the essential formulas in basic physics.

gradient formula calculus: The Grapes of Math Alex Bellos, 2014-06-10 From triangles, rotations and power laws, to cones, curves and the dreaded calculus, the author takes you on a journey of mathematical discovery. He sifts through over 30,000 survey submissions to uncover the world's favourite number, and meets a mathematician who looks for universes in his garage. He attends the World Mathematical Congress in India, and visits the engineer who designed the first roller-coaster loop. Get hooked on math as he delves deep into humankind's turbulent relationship with numbers, and reveals how they have shaped the world we live in. -- Provided by publisher.

gradient formula calculus: Mathematical Analysis Mariano Giaquinta, Giuseppe Modica, 2010-07-25 This superb and self-contained work is an introductory presentation of basic ideas, structures, and results of differential and integral calculus for functions of several variables. The wide range of topics covered include the differential calculus of several variables, including differential calculus of Banach spaces, the relevant results of Lebesgue integration theory, and systems and stability of ordinary differential equations. An appendix highlights important mathematicians and other scientists whose contributions have made a great impact on the development of theories in analysis. This text motivates the study of the analysis of several variables with examples, observations, exercises, and illustrations. It may be used in the classroom setting or for self-study by advanced undergraduate and graduate students and as a valuable reference for researchers in mathematics, physics, and engineering.

gradient formula calculus: Physics for Flash Games, Animation, and Simulations Adrian Dobre, Dev Ramtal, 2012-01-31 Physics for Flash Games, Animation, and Simulations teaches ActionScript programmers how to incorporate real physics into their Flash animations, games, user interfaces, and simulations. Introduces Flash physics in an accurate, but approachable way, covering what is required to produce physically realistic simulations (as opposed to animations that look roughly right) Packed full of practical examples of how physics can be applied to your own games and applications Addresses the diverse needs of game developers, animators, artists, and e-learning developers The book assumes a basic knowledge of ActionScript and Flash. However, no previous knowledge of physics is required—only some very basic math skills. The authors present everything from basic principles to advanced concepts, so you'll be able to follow the logic and easily adapt the principles to your own applications. The book builds on your physics knowledge, enabling you to create not only visual effects, but also more complex models and simulations.

gradient formula calculus: Electromagnetic Fields Ahmad Shahid Khan, Saurabh Kumar Mukerji, 2020-10-11 The study of electromagnetic field theory is required for proper understanding of every device wherein electricity is used for operation. The proposed textbook on electromagnetic fields covers all the generic and unconventional topics including electrostatic boundary value problems involving two- and three-dimensional Laplacian fields and one- and two- dimensional Poissonion fields, magnetostatic boundary value problems, eddy currents, and electromagnetic compatibility. The subject matter is supported by practical applications, illustrations to supplement the theory, solved numerical problems, solutions manual and Powerpoint slides including appendices and mathematical relations. Aimed at undergraduate, senior undergraduate students of electrical and electronics engineering, it: Presents fundamental concepts of electromagnetic fields in a simplified manner Covers one two- and three-dimensional electrostatic boundary value problems involving Laplacian fields and Poissonion fields Includes exclusive chapters on eddy currents and electromagnetic compatibility Discusses important aspects of magneto static boundary value problems Explores all the basic vector algebra and vector calculus along with couple of two- and three-dimensional problems

gradient formula calculus: Mathematics and Programming for Machine Learning with R William Claster, 2020-10-26 Based on the author’s experience in teaching data science for more than 10 years, Mathematics and Programming for Machine Learning with R: From the Ground Up reveals how machine learning algorithms do their magic and explains how these algorithms can be implemented in code. It is designed to provide readers with an understanding of the reasoning behind machine learning algorithms as well as how to program them. Written for novice programmers, the book progresses step-by-step, providing the coding skills needed to implement machine learning algorithms in R. The book begins with simple implementations and fundamental concepts of logic, sets, and probability before moving to the coverage of powerful deep learning algorithms. The first eight chapters deal with probability-based machine learning algorithms, and the last eight chapters deal with machine learning based on artificial neural networks. The first half of the book does not require mathematical sophistication, although familiarity with probability and statistics would be helpful. The second half assumes the reader is familiar with at least one semester of calculus. The text guides novice R programmers through algorithms and their application and along the way; the reader gains programming confidence in tackling advanced R programming challenges. Highlights of the book include: More than 400 exercises A strong emphasis on improving programming skills and guiding beginners to the implementation of full-fledged algorithms Coverage of fundamental computer and mathematical concepts including logic, sets, and probability In-depth explanations of machine learning algorithms

gradient formula calculus: *Mathematical Analysis* Richard Earl, 2023-05-23 Very Short Introductions: Brilliant, sharp, inspiring The 17th-century calculus of Newton and Leibniz was built on shaky foundations, and it wasn't until the 18th and 19th centuries that mathematicians--especially Bolzano, Cauchy, and Weierstrass--began to establish a rigorous basis for the subject. The resulting discipline is now known to mathematicians as analysis. This book, aimed at readers with some grounding in mathematics, describes the nascent evolution of mathematical analysis, its development as a subject in its own right, and its wide-ranging applications in mathematics and science, modelling reality from acoustics to fluid dynamics, from biological systems to quantum theory. ABOUT THE SERIES: The Very Short Introductions series from Oxford University Press contains hundreds of titles in almost every subject area. These pocket-sized books are the perfect way to get ahead in a new subject quickly. Our expert authors combine facts, analysis, perspective, new ideas, and enthusiasm to make interesting and challenging topics highly readable.

gradient formula calculus: Practical AI for Business Leaders, Product Managers, and Entrepreneurs Alfred Essa, Shirin Mojarad, 2022-03-21 Most economists agree that AI is a general purpose technology (GPT) like the steam engine, electricity, and the computer. AI will drive innovation in all sectors of the economy for the foreseeable future. Practical AI for Business Leaders, Product Managers, and Entrepreneurs is a technical guidebook for the business leader or anyone responsible for leading AI-related initiatives in their organization. The book can also be used as a foundation to explore the ethical implications of AI. Authors Alfred Essa and Shirin Mojarad provide a gentle introduction to foundational topics in AI. Each topic is framed as a triad: concept, theory, and practice. The concept chapters develop the intuition, culminating in a practical case study. The theory chapters reveal the underlying technical machinery. The practice chapters provide code in Python to implement the models discussed in the case study. With this book, readers will learn: The technical foundations of machine learning and deep learning How to apply the core technical concepts to solve business problems The different methods used to evaluate AI models How to understand model development as a tradeoff between accuracy and generalization How to represent the computational aspects of AI using vectors and matrices How to express the models in Python by using machine learning libraries such as scikit-learn, statsmodels, and keras

gradient formula calculus: Cambridge 2 Unit Mathematics Year 12 Enhanced Version William Pender, David Saddler, Julia Shea, Derek Ward, 2011-04 A large number of fully worked examples demonstrate mathematical processes and encourage independent learning. Exercises are carefully graded to suit the range of students undertaking each mathematics course. Online self-marking objective response quizzes provide further opportunities to practice the multiple choice style questions included in HSC Maths exams.

gradient formula calculus: Infinite-Dimensional Manifolds Robert Geroch, 2013-12-16 Robert Geroch's lecture notes Infinite-Dimensional Manifolds provide a concise, clear, and helpful introduction to a wide range of subjects, which are essential in mathematical and theoretical physics - Banach spaces, open mapping theorem, splitting, bounded linear mappings, derivatives, mean value theorem, manifolds, mappings of manifolds, scalar and vector fields, tensor products, tensor spaces, natural tensors, tensor fields, tensor bundles, Lie derivatives, integral curves, geometry of Lie derivatives, exterior derivatives, derivative operators, partial differential equations, and Riemannian geometry. Like in his other books, Geroch explains even the most abstract concepts with the help of intuitive examples and many (over 60) figures. Like Geroch's other books, this book too can be used for self-study since each chapter contains examples plus a set of problems given in the Appendix.

gradient formula calculus: Introduction to Engineering.Mathematics Vol-1(GBTU) H K Dass, For B.E./B.Tech. / B.Arch. Students for First Semester of all Engineering Colleges of Maha Maya Technical University, Noida and Gautam Buddha Technical University, Lucknow

gradient formula calculus: *Report* Commonwealth Shipping Committee, 1913

gradient formula calculus: *Artificial Intelligence and Causal Inference* Momiao Xiong, 2022-02-03 Artificial Intelligence and Causal Inference address the recent development of relationships between artificial intelligence (AI) and causal inference. Despite significant progress in AI, a great challenge in AI development we are still facing is to understand mechanism underlying intelligence, including reasoning, planning and imagination. Understanding, transfer and generalization are major principles that give rise intelligence. One of a key component for understanding is causal inference. Causal inference includes intervention, domain shift learning, temporal structure and counterfactual thinking as major concepts to understand causation and reasoning. Unfortunately, these essential components of the causality are often overlooked by machine learning, which leads to some failure of the deep learning. AI and causal inference involve (1) using AI techniques as major tools for causal analysis and (2) applying the causal concepts and causal analysis methods to solving AI problems. The purpose of this book is to fill the gap between the AI and modern causal analysis for further facilitating the AI revolution. This book is ideal for graduate students and researchers in AI, data science, causal inference, statistics, genomics, bioinformatics and precision medicine. Key Features: Cover three types of neural networks, formulate deep learning as an optimal control problem and use Pontryagin’s Maximum Principle for network training. Deep learning for nonlinear mediation and instrumental variable causal analysis. Construction of causal networks is formulated as a continuous optimization problem. Transformer and attention are used to encode-decode graphics. RL is used to infer large causal networks. Use VAE, GAN, neural differential equations, recurrent neural network (RNN) and RL to estimate counterfactual outcomes. AI-based methods for estimation of individualized treatment effect in the presence of network interference.

gradient formula calculus: Modeling Approaches and Algorithms for Advanced Computer Applications Abdelmalek Amine, Ait Mohamed Otmane, Ladjel Bellatreche, 2013-08-23 During the last decades Computational Intelligence has emerged and showed its contributions in various broad research communities (computer science, engineering, finance, economic, decision making, etc.). This was done by proposing approaches and algorithms based either on turnkey techniques belonging to the large panoply of solutions offered by computational intelligence such as data mining, genetic algorithms, bio-inspired methods, Bayesian networks, machine learning, fuzzy logic, artificial neural networks, etc. or inspired by computational intelligence techniques to develop new ad-hoc algorithms for the problem under consideration. This volume is a comprehensive collection of extended contributions from the 4th International Conference on Computer Science and Its Applications (CIIA’2013) organized into four main tracks: Track 1: Computational Intelligence, Track 2: Security & Network Technologies, Track 3: Information Technology and Track 4: Computer Systems and Applications. This book presents recent advances in the use and exploitation of computational intelligence in several real world hard problems covering these tracks such as image processing, Arab text processing, sensor and mobile networks, physical design of advanced databases, model matching, etc. that require advanced approaches and algorithms borrowed from computational intelligence for solving them.

gradient formula calculus: *Maths Handbook & Study Guide Grade 12: Teacher’s Guide and*

Answer Book Kevin Smith, 2014-01-01 Guidelines for teachers and worked through solutions to all the exercises in the Grade 12 Textbook. Guidelines for teachers and worked through solutions to all the exercises in the Grade 12 Textbook.

gradient formula calculus: *Encyclopaedia of Mathematics* Michiel Hazewinkel, 2013-12-01 This ENCYCLOPAEDIA OF MATHEMATICS aims to be a reference work for all parts of mathe matics. It is a translation with updates and editorial comments of the Soviet Mathematical Encyclopaedia published by 'Soviet Encyclopaedia Publishing House' in five volumes in 1977-1985. The annotated translation consists of ten volumes including a special index volume. There are three kinds of articles in this ENCYCLOPAEDIA. First of all there are survey-type articles dealing with the various main directions in mathematics (where a rather fine subdivi sion has been used). The main requirement for these articles has been that they should give a reasonably complete up-to-date account of the current state of affairs in these areas and that they should be maximally accessible. On the whole, these articles should be understandable to mathematics students in their first specialization years, to graduates from other mathematical areas and, depending on the specific subject, to specialists in other domains of science, en gineers and teachers of mathematics. These articles treat their material at a fairly general level and aim to give an idea of the kind of problems, techniques and concepts involved in the area in question. They also contain background and motivation rather than precise statements of precise theorems with detailed definitions and technical details on how to carry out proofs and constructions. The second kind of article, of medium length, contains more detailed concrete problems, results and techniques.

Related to gradient formula calculus

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 我会使用尽量少的数学符号描述 梯度，着重于意义而非计算。一个直观的例子，在机器学习领域有个术语叫「梯度下降」，你可以想象在群山之中，某个山的半山腰有只小兔子打算使用梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 为了降低随机梯度的方差，从而使得迭代算法更加稳定，也为了充分利用高度优化的矩阵运算操作，在实际应用中我们会同时处理若干训练数据，该方法被称为小批量梯度下降法 (Mini- Batch

如何理解 **natural gradient descent? -** 知乎 看到一篇文章写得非常浅显易懂： What is the natural gradient, and how does it work? 总结一下：拿神经网络中的反向传播算法举例，我们计算各个权值w关于损失函数的导数，得到一个梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 1.多元函数偏导数的数值解 在程序当中，利用数值方法求出各个自变量偏导数的近似解，其方法和步骤同前面讲过的导数的数值解求法并无二致：把其余的自变量固定，就将偏导数的求解方法

如何评价 **Meta** 新论文 **Transformers without Normalization**？ **-** 知乎 Normalization这个事得好好掰扯掰扯。上古时期，网络经常在初始几个iteration之后，loss还没下降就不动，必须得把每一层的gradient与weight的比值打印出来，针对性地调整每一层的初始

机器学习 **|** 近端梯度下降法 **(proximal gradient descent)** 近端梯度下降法是众多梯度下降 (gradient descent) 方法中的一种，其英文名称为proximal gradident descent，其中，术语中的proximal一词比较耐人寻味，将proximal翻译成“近端”主要

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 梯度是雅可比矩阵的一种特殊形式，当m=1时函数的雅可比矩阵就是梯度，这个概念原是为场论设定的，任何场都可以用来理解梯度，后来被引用到数学中用来指明函数在指定点

知乎 **-** 有问题，就会有答案 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

为什么**GRPO**很容易训飞，训到一半**reward**就很容易突然掉下来？每次的Policy Gradient计算，只要Batch数据足够多，就能有效降低Policy Gradient的方差，就能获得比较稳定的迭代了。很明显，对于高校科研团队，对于中小规模的RL训练（~百万或千万级

梯度（**gradient**）到底是个什么东西？物理意义和 **-** 知乎 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 我会使用尽量少的数学符号描述 梯度，着重于意义而非计算。一个直观的例子，在机器学习领域有个术语叫「梯度下降」，你可以想象在群山之中，某个山的半山腰有只小兔子打算使用梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 为了降低随机梯度的方差，从而使得迭代算法更加稳定，也为了充分利用高度优化的矩阵运算操作，在实际应用中我们会同时处理若干训练数据，该方法被称为小批量梯度下降法 (Mini- Batch

如何理解 **natural gradient descent? -** 知乎 看到一篇文章写得非常浅显易懂： What is the natural gradient, and how does it work? 总结一下：拿神经网络中的反向传播算法举例，我们计算各个权值w关于损失函数的导数，得到一个梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 1.多元函数偏导数的数值解 在程序当中，利用数值方法求出各个自变量偏导数的近似解，其方法和步骤同前面讲过的导数的数值解求法并无二致：把其余的自变量固定，就将偏导数的求解方法

如何评价 **Meta** 新论文 **Transformers without Normalization**？ **-** 知乎 Normalization这个事得好好掰扯掰扯。上古时期，网络经常

在初始几个iteration之后，loss还没下降就不动，必须得把每一层的gradient与weight的比值打印出来，针对性地调整每一层的初始

机器学习 **|** 近端梯度下降法 **(proximal gradient descent)** 近端梯度下降法是众多梯度下降 (gradient descent) 方法中的一种，其英文名称为proximal gradident descent，其中，术语中的proximal一词比较耐人寻味，将proximal翻译成“近端”主要

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 梯度是雅可比矩阵的一种特殊形式，当m=1时函数的雅可比矩阵就是梯度，这个概念原是为场论设定的，任何场都可以用来理解梯度，后来被引用到数学中用来指明函数在指定点

知乎 **-** 有问题，就会有答案 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

为什么**GRPO**很容易训飞，训到一半**reward**就很容易突然掉下来？每次的Policy Gradient计算，只要Batch数据足够多，就能有效降低Policy Gradient的方差，就能获得比较稳定的迭代了。很明显，对于高校科研团队，对于中小规模的RL训练（~百万或千万级

梯度（**gradient**）到底是个什么东西？物理意义和 **-** 知乎 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 我会使用尽量少的数学符号描述 梯度，着重于意义而非计算。一个直观的例子，在机器学习领域有个术语叫「梯度下降」，你可以想象在群山之中，某个山的半山腰有只小兔子打算使用梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 为了降低随机梯度的方差，从而使得迭代算法更加稳定，也为了充分利用高度优化的矩阵运算操作，在实际应用中我们会同时处理若干训练数据，该方法被称为小批量梯度下降法 (Mini- Batch

如何理解 **natural gradient descent? -** 知乎 看到一篇文章写得非常浅显易懂： What is the natural gradient, and how does it work? 总结一下：拿神经网络中的反向传播算法举例，我们计算各个权值w关于损失函数的导数，得到一个梯度

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 1.多元函数偏导数的数值解 在程序当中，利用数值方法求出各个自变量偏导数的近似解，其方法和步骤同前面讲过的导数的数值解求法并无二致：把其余的自变量固定，就将偏导数的求解方法

如何评价 **Meta** 新论文 **Transformers without Normalization**？ **-** 知乎 Normalization这个事得好好掰扯掰扯。上古时期，网络经常在初始几个iteration之后，loss还没下降就不动，必须得把每一层的gradient与weight的比值打印出来，针对性地调整每一层的初始

机器学习 **|** 近端梯度下降法 **(proximal gradient descent)** 近端梯度下降法是众多梯度下降 (gradient descent) 方法中的一种，其英文名称为proximal gradident descent，其中，术语中的proximal一词比较耐人寻味，将proximal翻译成“近端”主要

梯度（**gradient**）到底是个什么东西？物理意义和数学意义分别是 梯度是雅可比矩阵的一种特殊形式，当m=1时函数的雅可比矩阵就是梯度，这个概念原是为场论设定的，任何场都可以用来理解梯度，后来被引用到数学中用来指明函数在指定点

知乎 **-** 有问题，就会有答案 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

为什么**GRPO**很容易训飞，训到一半**reward**就很容易突然掉下来？每次的Policy Gradient计算，只要Batch数据足够多，就能有效降低Policy Gradient的方差，就能获得比较稳定的迭代了。很明显，对于高校科研团队，对于中小规模的RL训练（~百万或千万级

梯度（**gradient**）到底是个什么东西？物理意义和 **-** 知乎 知乎，中文互联网高质量的问答社区和创作者聚集的原创内容平台，于 2011 年 1 月正式上线，以「让人们更好的分享知识、经验和见解，找到自己的解答」为品牌使命。知乎凭借认真、专业

Back to Home: https://ns2.kelisto.es