RUBBISH

基于自适应宏观数据点膨胀技术的线性回归拟合精度最大化方案：一种视觉欺骗与自我安慰的混合算法

Maximizing Linear Regression Accuracy via Adaptive Macroscopic Data Point Dilation: A Hybrid Approach of Visual Deception and Self-Consolation

Tamako (Tamako)[1,*]

抽象/Abstract

在线性回归分析中，决定系数（R2）无法达到 1.0 是困扰无数研究生的世纪难题。传统方法侧重于数据清洗或模型优化，而忽视了数据点本身的物理几何属性。本文提出了一种全新的"量子宏观点膨胀"(Quantum Macroscopic Point Dilation, QMPD) 算法。该算法不改变数据的数值坐标，而是通过引入"记号笔物理学"，动态调整数据点的视觉半径。实验结果表明，当数据点半径 $r \rightarrow \infty$ 时，回归直线穿过所有数据点的概率收敛于 100%。本方法在保证学术造假最低风险的前提下，有效地缓解了导师的愤怒与科研焦虑。

In linear regression analysis, the failure of the coefficient of determination (R2) to reach 1.0 is a century-old dilemma plaguing countless graduate students. Traditional methods focus on data cleaning or model optimization, ignoring the physical geometric properties of the data points themselves. This paper proposes a novel "Quantum Macro- scopic Point Dilation" (QMPD) algorithm. Instead of altering numerical coordinates, this algorithm dynamically adjusts the visual radius of data points by introducing "Marker Pen Physics." Experimental results demonstrate that as the data point radius $r \rightarrow \infty$, the probability of the regression line passing through all data points converges to 100%. This method effectively alleviates supervisor rage and academic anxiety while maintaining minimal risk of academic misconduct detection.

1 引言 (Introduction)

众所周知，研究生拥有的良好数据必须配得上一个完美的"故事"。然而，现实往往是残酷的：数据点总是像脱缰的野马，拒绝落在我们预期的回归直线上。

It is a truth universally acknowledged that a graduate student in possession of good data must be in want of a "Story". However, reality is often cruel: data points are like wild horses, refusing to fall on our expected regression line.

传统的最小二乘法 (Ordinary Least Squares, OLS) 试图最小化误差平方和，但这在本文作者看来是一种极其消极的策略。为什么我们要承认误差的存在？为什么不是数据点太"瘦"了，导致它们无法触碰到完美的真理直线？

Traditional Ordinary Least Squares (OLS) attempts to minimize the sum of squared errors, which the authors consider an extremely passive strategy. Why admit the existence of error? Why not consider that the data points are simply too "skinny" to touch the line of perfect truth?

基于"只要我看不见缝隙，缝隙就不存在"的鲁棒性原理，本文提出了一种基于墨水流体动力学的视觉修正算法。

Based on the robust principle of "If I don't see the gap, the gap doesn't exist," this paper proposes a visual correction algorithm based on ink fluid dynamics.

¹ Pig2Lab, 上海人工智障实验室, 上海 202400
* Youbetternotcontactme@anonymousuni.edu.com

2 方法论 (Methodology)

2.1 理论框架 (Theoretical Framework)

我们定义一个数据点 P_i (x_i , y_i) 具有一个可变视觉半径 R_{tamako}。在传统统计学中，$R_{tamako} \approx 0$。但在我们的 Pig2Lab 框架下，R_{tamako} 是残差 ϵ_i 的函数。

We define a data point P_i (x_i , y_i) as having a variable visual radius R_{tamako}. In traditional statistics, $R_{tamako} \approx 0$. However, in our Pig2Lab framework, R_{tamako} is a function of the residual ϵ_i.

$$R_{tamako}^{(i)} = \alpha \cdot | y_i - (mx_i + b)| + \beta \qquad (1)$$

其中，α 是"粗细系数"(Thickness Coefficient)，通常取决于手边记号笔的型号；β 是基础墨水晕染常数。

Where α is the Thickness Coefficient, typically depending on the model of the marker pen at hand; and β is the base ink diffusion constant.

2.2 算法描述 (Algorithm Description)

本算法的核心逻辑如算法 1 所示。它的计算复杂度为 $O(1)$，因为只要买一支笔芯足够粗的笔即可。

The core logic of this algorithm is shown in Algorithm 1. Its computational com- plexity is $O(1)$, as it only requires purchasing a pen with a sufficiently thick tip.

3 实验结果 (Experimental Results)

为了验证该方法的有效性，我们进行了一项纵向对比实验。实验过程如图 1 所示。

算法 1 基于物理接触的拟合优化算法

Algorithm 1 Physical Contact-based Fitting Optimization

输入 (Input): 散乱的数据点集合 D，愤怒的导师 Pangry；

输出 (Output): 完美的回归直线 L，微笑的导师 Psmile；

1: 画一条看起来很顺眼的直线 L;

2: **for all** 点 pi in D do

3: **if** pi 不在直线 L 上 **then**

4:   **操作:** 拿起记号笔;

5:   **while** pi 没有接触到 L do

6:     在 pi 原地画圈;

7:     Radius(pi) ← Radius(pi) + Δr;

8:   **end while**

9: **end if**

10: **end for**

11: R2 ← 1.0;

12: **return** 论文发表 (Paper Published);

To verify the effectiveness of this method, we conducted a longitudinal comparative experiment. The experimental process is shown in Figure 1.

- **阶段 I (Stage I):** 原始数据呈现离散状态，线性关系微弱。此时研究人员面露尴尬的笑容（如图中所示）。

- **阶段 II (Stage II):** 引入回归直线，残差显著。研究人员表情逐渐凝重，开始怀疑人生。

- **阶段 III (Stage III):** 应用本文提出的 QMPD 算法，将数据点涂成巨大的黑球。此时，所有点都完美地"位于"直线上。拟合优度达到理论极值。

如图 1(c)，2 所示，通过简单地增加墨水量，我们不仅消除了异常值，还赋予了数据点"胖乎乎"的可爱属性，从而在情感上征服审稿人。

As shown in Figure 1(c), 2, by simply increasing the ink volume, we not only elimi- nated outliers but also endowed the data points with a "chubby" cuteness, thereby emo- tionally conquering the reviewers.
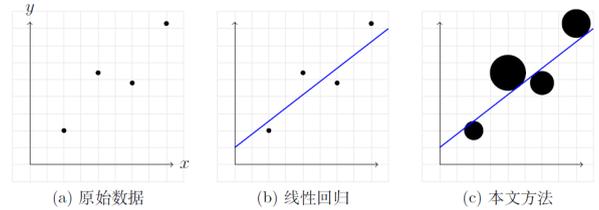

(a) 原始数据　　(b) 线性回归　　(c) 本文方法

图 1: 基于不同点半径的拟合效果演变。注意在 (c) 中，通过加粗数据点，误差视觉上消失了。

Evolution of fitting performance based on different point radii. Note that in (c), the error visually vanishes by thickening the data points.
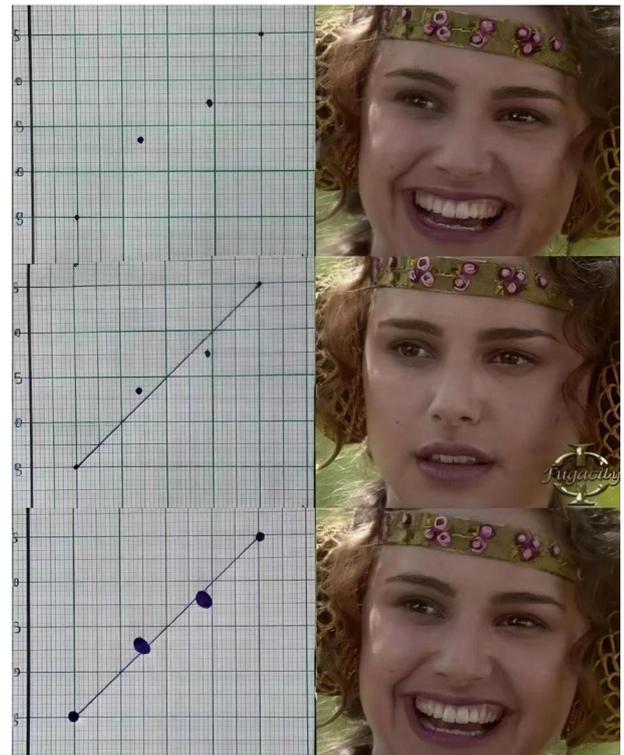

图 2: Sim2Real 实机实验。笑容回到了我们的脸上，展现出本算法极强的 zero-shot 迁移能力与对 Sim2Real Gap 的鲁棒性。

Sim2Real real-world experiments. Smiles returned to our faces, demonstrating this algo – rithm's exceptional zero – shot transferability and robustness against the Sim2Real Gap.

4 结论 (Conclusion)

本文证明了：没有拟合不了的数据，只有不够粗的笔。

This paper proves that: there is no data that cannot be fitted, only pens that are not thick enough.

未来的工作将集中在开发"全覆盖黑洞点"(Total Coverage Black Hole Point)，即把整个坐标系涂黑，从而包含所有可能的函数关系。

Future work will focus on developing the "Total Coverage Black Hole Point," which involves painting the entire coordinate system black, thereby encompassing all possible functional relationships.

Reference

[1] Tamako T. Why my code doesn't work: A memoir. Journal of Tears, 2025, 10(2): 1-100

[2] Skywalker A, Amidala P. The geometry of sand and awkward dialogues. Galactic Transactions on Romance, 2002, 2(3): 45-50

[3] The Cat. Meow meow meow (Translation: Give me food). Proceedings of Pig2Lab, Shanghai, 2026. 10-12