

不稳定神经网络中的反向传播算法

ZHEYU XIE¹, ANGXIU NI^{2,3}

1. 摘要

在不稳定神经网络中，梯度爆炸问题限制了反向传播算法的有效性。随着网络层数和复杂度增加，梯度可能会指数级增长，导致训练过程中数值不稳定和模型性能下降。本文回顾了不稳定神经网络的理论基础，包括李雅普诺夫谱和李雅普诺夫向量的概念，用于描述系统的动态特性和稳定性。伴随李雅普诺夫谱和对偶性的概念对于解决梯度爆炸问题很重要。

传统反向传播算法中，梯度爆炸问题的解决方法包括梯度裁剪和正则化技术，但在不稳定神经网络中效果有限。为了克服这个挑战，本文提出了一种基于伴随阴影的新反向传播方法，利用伴随李雅普诺夫谱的信息来调整梯度的传播路径和强度，有效地缓解梯度爆炸问题。同时，介绍了核微分方法，通过引入核函数平滑梯度计算，提高了计算的稳定性和准确性。

本文在理论层面分析了传统反向传播算法在不稳定神经网络中的表现和局限性，强调了梯度爆炸问题对参数更新和模型训练的影响。基于伴随阴影的反向传播方法重新定义了梯度更新规则，并通过实验验证了其在不同类型不稳定神经网络中的有效性。实验结果表明，该方法显著减小梯度爆炸的影响，提升了模型的收敛速度和性能稳定性。

为了验证方法的广泛适用性，本文将核微分方法与伴随阴影技术相结合，构建了一种混合优化算法。实验结果显示，与传统方法相比，新的混合优化算法在训练速度、收敛性和最终模型性能方面有显著提升。这表明核微分方法在处理梯度爆炸问题时提供了额外的平滑效果，使得梯度更新过程更加稳定。

综上所述，本文通过理论分析和实验验证，提出了一种创新的解决不稳定神经网络中梯度爆炸问题的方法。基于伴随阴影的反向传播方法和核微分方法的结

¹ DEPARTMENT OF MATHEMATICS, TSINGHUA UNIVERSITY, BEIJING, CHINA.

² DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, IRVINE, USA

³ YAU MATHEMATICAL SCIENCES CENTER, TSINGHUA UNIVERSITY, BEIJING, CHINA.

E-mail address: xie.zeyu20@gmail.com, niangxiu@gmail.com.

Date: 2024 年 5 月 31 日.

合为未来研究和应用提供了新的方向和思路。这些研究结果不仅加深了对不稳定神经网络动态特性的理解，也为改进反向传播算法提供了新的工具和方法。

2. 绪论

2.1. 文献综述.

2.2. 现有成果.

3. 不稳定神经网络

3.1. 李雅普诺夫谱.

Definition 1 (李雅普诺夫常数). 设 A 是一个 $n \times n$ 的矩阵, x 是一个 n 维向量。如果存在一个实数 λ , 使得对于任意的 $x \neq 0$, 都有

$$\frac{\|Ax\|}{\|x\|} \leq \lambda$$

则称 λ 是矩阵 A 的李雅普诺夫常数, 记作 $\lambda(A)$ 。

Definition 2 (李雅普诺夫谱). 设 A 是一个 $n \times n$ 的矩阵, $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值。则 A 的李雅普诺夫谱定义为

$$\Lambda(A) = \max_{1 \leq i \leq n} |\lambda_i|$$

3.2. 李雅普诺夫向量.

Definition 3 (李雅普诺夫向量). 设 A 是一个 $n \times n$ 的矩阵, x 是一个 n 维向量。如果存在一个实数 λ , 使得对于任意的 $x \neq 0$, 都有

$$\frac{\|Ax\|}{\|x\|} \leq \lambda$$

3.3. 伴随李雅普诺夫谱和对偶性.

Definition 4 (伴随李雅普诺夫谱). 设 A 是一个 $n \times n$ 的矩阵, B 是一个 $n \times n$ 的矩阵。 A 的伴随李雅普诺夫谱定义为

$$\Lambda(A, B) = \max_{1 \leq i \leq n} |\lambda_i|$$

4. 梯度爆炸下的反向传播算法

首先介绍神经网络中反向传播算法的基本原理和流程。

反向传播算法用到了链式法则，通过逐层计算梯度来更新网络参数。在不稳定神经网络中，梯度爆炸问题会导致梯度的指数级增长，使得参数更新过程不稳定。传统的反向传播算法在处理梯度爆炸问题时效果有限，需要引入新的方法和技术来提高算法的稳定性和收敛性。

我们分别以两个例子来说明反向传播算法的具体实现。

【具体介绍两个例子，传统神经网络和 RNN】

4.1. 传统方法的困境.

4.2. 通过伴随阴影进行反向传播.

4.3. 核微分方法.

5. 致谢

在本科学习和论文撰写的过程中，我得到了许多人的帮助和支持。在此，我要向所有在我这段旅程中给予帮助和支持的人表示衷心的感谢。

首先，我要特别感谢我的导师倪昂修老师。倪老师不仅在本论文的选题、研究方法和数据分析等方面给予了我悉心的指导，还在我遇到困惑和困难时给予了极大的耐心和鼓励。他的深厚学术造诣和认真负责的态度对我产生了深远的影响，使我受益匪浅。

其次，我要感谢我的家人。你们无条件的支持和鼓励是我前进的不竭动力。在我遇到困难和挫折时，你们总是给予我最温暖的关怀，让我能够保持积极向上的心态。

此外，我还要感谢我的朋友、高中同学夏斐然，他在我学习和生活中给予了我很多帮助和支持。在我遇到困难和挫折时，他总是给予我最温暖的关怀，让我能够保持积极向上的心态。

最后，我也要感谢前来参与答辩的专家和评审老师。感谢您们抽出宝贵的时间来审阅我的论文，提出宝贵的意见和建议。您的指导和批评对我今后的学习和研究具有重要的指导意义。

再次感谢所有在我学术道路上给予帮助和支持的人。希望通过这篇论文，我能为机器学习、深度学习等领域的研究工作贡献自己的一份力量。