

# 自动微分与编程语言<sup>\*</sup>

刘金国<sup>1)</sup> 许开来<sup>2)</sup>

1) (哈佛大学物理系, 波士顿 02138)

2) (...)

自动微分是机器学习中核心的数值技术。机器学习中用到的自动微分主要是以张量层面的微分居多, 其衍生的工具可以解决张量网络等算法的变分优化。而本文主要介绍另外一种更加通用的, 在编程语言层面的自动微分, 并介绍它在科学计算中的用途和未来发展的趋势。

关键词: 自动微分, 科学计算, 编程语言

**PACS:** 02.60.Pn, 02.30.Jr

## 1 引言

自动微分是指自动获取一个目标计算机程序导数的技术。与很多人印象不同的是, 自动微分是个十分有年代的技术。Nolan 曾在他 1953 年的博士论文中就提出过 [Nolan(1953)] 自动化求导计算机程序的构想, 它利用了计算机程序的一个普遍特性, 那就是不管程序有多么复杂, 总可以被分解成加减乘除之类的基础指令。我们很容易知晓基础指令的自动微分规则, 因而可以通过简单的链式法则推导出整个程序的导数。在上世纪 60 年代, 自动微分发展出了两种不同的实践, 一种叫做前向传播 (Forward propagation) [Wengert(1964)], 一种叫做后向传播 (Backward propagation) [Boltysanski *et al.*(1960) Boltysanski, Gamkrelidze, Mi]. 一个计算过程  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$  可以被描述为

$$\mathbf{x}^1 = f_1(\mathbf{x}^0)$$

---

\*

† 通信作者. E-mail: cacate0129@gmail.com

$$\mathbf{x}^2 = f_2(\mathbf{x}^1)$$

...

$$\mathbf{x}^L = f_L(\mathbf{x}^{L-1})$$

其中  $x^0 \in R^m$ ,  $x^L \in R^n$ ,  $L$  是计算的步骤数。这段程序的雅可比 (Jacobian) 矩阵  $J_{ij} \equiv \frac{\partial x_i^L}{\partial x_j^0}$  是一个  $n \times m$  的矩阵, 其中  $x_j^0$  和  $x_i^L$  代表了输入和输出中的一个元素。前向自动微分的链式法则表述为  $\frac{\partial \mathbf{x}^k}{\partial x_j^0} = \frac{\partial \mathbf{x}^k}{\partial \mathbf{x}^{k-1}} \frac{\partial \mathbf{x}^{k-1}}{\partial x_j^0}$ , 它计算了雅可比矩阵的第  $j$  列。后向自动微分的链式法则为  $\frac{\partial \mathbf{x}_i^L}{\partial x^{k-1}} = \frac{\partial \mathbf{x}_i^L}{\partial \mathbf{x}^k} \frac{\partial \mathbf{x}^k}{\partial x^{k-1}}$ , 它计算了雅可比矩阵的第  $i$  行。不难发现, 它们梯度的传递计算方向分别是向前和向后, 这也决定了它们的实现难度。

过去, 很多机器学习中的神经网络往往由简单的几种定义在张量上的操作构成, 因此人们很容易手推出一个常用函数族的导数。但最近几年, 机器学习界风靡起了微分编程。因为大家认识到神经网络的固定的函数族 (矩阵乘法, 卷积函数, relu 函数等) 已经开始渐渐满足不了神经网络多样性的设计, 我们需要把自动微分当作“一等公民”的编程语言。类似的需求在物理中也很多见, 比如在张量网络等模拟中需要用到的复数奇异值分解函数 [Wan and Zhang(2019), Liao *et al.*(2019)Liao, Liu, Wang, and Xiang], 严格对教化求解基态过程中用到的极大极小本正值求解器 [Hao Xie and Wang()], 模拟变分量子算法中么正量子线路的求导 [Luo *et al.*(2019)Luo, Liu, Zhang, and Wang]。人们为了手推这些求导规则付出了很多的努力, 如果我们有了编程语言层面的自动微分, 那么这些都可以在程序编写之时就得到解决。

## 2 自动微分在科学计算领域的应用

### 2.1 地震波的模拟

### 2.2 热带 (Tropical) 张量网络求解自旋玻璃最优构型

热带张量网络是张量网络的一个特别的应用, 它重新定义了张量中基础元素的代数

$$x \oplus y = \max(x, y), \quad x \odot y = x + y. \quad (1)$$

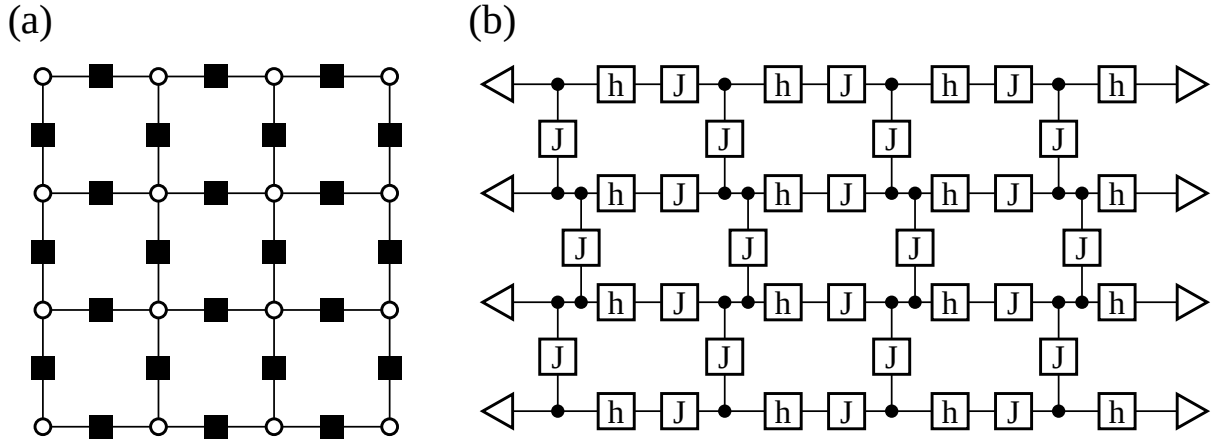


图 1: 正方晶格上的自旋玻璃问题, (a) 对应的热带张量网络表示。(b) 对应的“量子线路”表示。如果我们将张量中的元素定义为自旋玻璃的耦合强度, 我们就可以通过收缩这个张量结构得到自旋玻璃的最优构型。这个最优构型的能量表达式为

$$E(\{\sigma\}) = \max_{\sigma} - \sum_{i \sim j} J_{ij} \sigma_i \sigma_j - \sum h_i \sigma_i, \quad (2)$$

因此对它的微分就是。但是作了这样的替换, 张量收缩的微分规则就发生了变换。

### 3 自动微分与可逆计算

自动微分

### 4 结 论

### 致谢

感谢北京大学力学系某某教授和某某博士以及某某的讨论.

### 附录 A1

标题排列和编号方式为 A1, A2, A3.

## 附录 A2

- [Nolan(1953)] J. F. Nolan, *Analytical differentiation on a digital computer*, Ph.D. thesis, Massachusetts Institute of Technology (1953).
- [Wengert(1964)] R. E. Wengert, Communications of the ACM **7**, 463 (1964).
- [Boltyanski *et al.*(1960)]Boltyanski, Gamkrelidze, Mishchenko, and Pontryagin] V. Boltyanski, R. Gamkrelidze, E. Mishchenko, and L. Pontryagin, IFAC Proceedings Volumes **1**, 464 (1960).
- [Wan and Zhang(2019)] Z.-Q. Wan and S.-X. Zhang, “Automatic differentiation for complex valued svd,” (2019), arXiv:1909.02659 [math.NA] .
- [Liao *et al.*(2019)]Liao, Liu, Wang, and Xiang] H.-J. Liao, J.-G. Liu, L. Wang, and T. Xiang, \bibfield journal \bibinfo journal Physical Review X\ \textbf{\bibinfo volume 9} (\bibinfo year 2019),\ 10.1103/physrevx.9.031041.
- [Hao Xie and Wang()] J.-G. L. Hao Xie and L. Wang, \ arXiv:2001.04121 .
- [Luo *et al.*(2019)]Luo, Liu, Zhang, and Wang] X.-Z. Luo, J.-G. Liu, P. Zhang, and L. Wang, “Yao.jl: Extensible, efficient framework for quantum algorithm design,” (2019), arXiv:1912.10877 [quant-ph] .

# English Title \*

Guan Yun-Chang<sup>1)2)</sup>    Liu Bei<sup>1)†</sup>    Zhuge Liang<sup>2)</sup>

1) ( *State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Opto-Electronics, Shanxi University, Taiyuan 030006, China* )

2) ( *Department of Physics, Tsinghua University, Beijing 100084, China* )

## Abstract

To determine the probe made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. The sequence of amino acids in a protein is defined by a gene and encoded in the genetic code. This can happen either before the protein is used in the cell, or as part of control mechanisms.

**Key-**        Keyword1, Keyword2, Keyword3, Keyword4

**PAICS:** 02.10.Yn, 33.15.Vb, 98.52.Cf, 78.47.dc

---

\* Project supported by the State Key Development Program for Basic Research of China (Grant No. 2011CB00000), the National Natural Science Foundation of China (Grant Nos. 123456, 567890), and the National High Technology Research and Development Program of China (Grant No. 2011AA06Z000).

† Corresponding author. E-mail: aaa@bbb.ccc