

自动微分以及它在物理模拟中的应用^{*}

刘金国¹⁾ 许开来²⁾

1) (哈佛大学物理系, 坎布里奇 02138)

2) (斯坦福大学, 斯坦福 94305)

自动微分是利用计算机自动化求导的技术, 最近几十年因为被用于机器学习研究而被很多人了解。如今越来越多的物理学者意识到高效, 高效的, 自动化的求导可以对很多物理问题的求解提供新的思路。其中自动微分在物理模拟问题中的应用尤为重要且具有挑战性。本文介绍如何将自动微分技术运用到物理模拟的求导中, 介绍了共轭态法, 前向自动微分, 后向自动微分以及可逆计算自动微分的基本概念, 并横向对比了它们在物理模拟中的优势和劣势。

关键词: 自动微分, 科学计算, 可逆计算, Treeverse, 物理模拟

PACS: 02.60.Pn, 02.30.Jr, 91.30.-f

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。很多人了解它是因为它在机器学习中的成功应用, 人们可以用它优化带有千亿参数的神经网络 [Rosset(2019)]。与很多人印象不同的是, 自动微分其实是个很古老的技术。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, 而最近十几年, 自动微分在科学计算中的应用越来越广泛并拓展了人们可解决问题的范畴。一个典型的例子是变分蒙特卡洛算法。过去, 人们会将变分基态假设为 Gutzwiller 波函数 [Gutzwiller(1963)] 这样具有良好解析性质的函数。直到 2017 年, Carleo 等人 [Carleo and Troyer(2017), Deng *et al.*(2017)] Deng, Li, and Das Sa

*

† 通信作者. E-mail: jinguoliu@g.harvard.edu

把机器学习的一些变分模型限制波尔兹曼机 (RBM) 带入到了大家的视野。由于 RBM 的解析性质很好, 计算所需的求导过程可以通过解析推导, 所以当时人们并没有强烈的利用计算机辅助求导的动机。但受此启发, 大家意识到不光是 RBM 这样解析性质好的函数, 任何机器学习中的神经网络也可以被用做波函数的猜测形式 [Cai and Liu(2018)]。而且如果可以用流行的机器学习框架, 比如脸书公司开发的 PyTorch 和谷歌公司开发的 TensorFlow 来编写生成波函数的代码, 人们可以避免手动推导导数的麻烦, 从而可以把波函数的表达变得更加自由。而现在, 随着人们对微分的认知更加深刻, 波函数不光可以表达为以张量运算为主体的神经网络, 还可以是几乎任意的计算机程序。除了这个例子, 科学家们利用方便的, 自动化的计算机辅助求导还解决了很多包括量子模拟 [Luo *et al.*(2019)Luo, Liu, Zhang, and Wang], 张量网络 [Liao *et al.*(2019)Liao, Liu, Wang, and Xiang], 组合优化 [Liu *et al.*(2020)Liu, Wang, and Zhang] 等领域中的问题。甚至是一些非解析的蒙特卡洛抽样过程, 人们也设计出了一些办法对其自动化的求导 [Zhang *et al.*(2019)Zhang, Liu, Wang, and Xiang]。

虽然有这么多成功的例子, 但当人们把自动微分技术应用到物理模拟过程, 比如电路仿真, 海洋学 [Heimbach *et al.*(2019)Heimbach, Wang, and Zhang] 和地震学 [Symes(2007), Zhu *et al.*(2020)Zhu, Xu, Darve, and Beroza] 等, 常见的机器学习库经常无法直接胜任。一方面现有的机器学习框架中的自动微分需要存储程序每一步计算的部分状态以便在后向传播过程中取出用于计算局域梯度。而另一方面, 物理模拟过程经常包含物理量对时间的积分, 积分步长很小而步骤数很多, 导致记录中间状态对空间的需求巨大。为了解决自动微分对空间的需求, 人们可以假设在较短时间内积分器可逆来回溯状态, 也被称为共轭态法 [Plessix(2006), Chen *et al.*(2018)Chen, Rubanova, Bettencourt, and Duvenaud]。事实上, 除了 Leap frog 积分器在时间反演不变的哈密顿量问题中可以做到严格时间反演对称, 大多数问题中的积分器并不能保证可逆性, 所以共轭态法往往存在一定的由积分器带来的系统性误差。后来, 有人把机器学习中的 Treeverse 算法带入到了物理模拟的状态回溯中, 使得自动微分不再有积分器不可逆带来的系统误差。同时, Treeverse 算法可以在对数时间和空间下做到状态的严格回溯。本文提出了第三种节省空间的方案, 那就是利用可逆编程的方案利用 Bennett 算法来进行状态回溯, 并横向对比了共轭态 (Adjoint-State) 方法, 前向自动微分以及基于 Treeverse 算法和可逆计算的后向自动微分在处理物理模拟问题中的优劣。

章节 2 介绍了共轭态法和自动微分的基本原理。章节 3 介绍了基于检查点和可逆编程的两种后向自动微分的基础理论, 尤其两者如何权衡程序的运行时间和空间。章节 4 介绍了不同自动微分技术在地震波模拟过程中的应用。

2 对物理模拟的自动微分方法

物理模拟过程的常见求解方案是将偏微分方程的空间部分离散并作差分处理 [Grote and Sim(2010)], 将其转换为一个对时间的常微分方程

$$\frac{ds}{dt} = f(s, t, \theta)$$

其中 s 为状态, t 为时间, θ 为控制参数。假设我们已经拥有一个常微分方程求解器来得到末态

$$\begin{aligned} s_n &= \text{ODESolve}(s_0, f, t_0, t_n, \theta) \\ &= (s_{i+1} = \text{ODEStep}(s_i, t_i, \theta, \Delta t) \text{ for } i = 1, 2, \dots, n) \end{aligned} \quad (1)$$

其中 t_0 和 s_0 分别为初始时间和状态, $t_n = t_0 + n\Delta t$ 和 s_n 为末了时间和状态。这个常微分方程求解器在求解过程中会把时间离散化, 作 n 步叠代, 每步仅做从时刻 t_i 到时刻 $t_i + \Delta t$ 的演化。最后我们还可以定义一个损失函数 $\mathcal{L} = \text{loss}(s_n)$ 。自动微分的目标则是求解损失量对参数的导数 $\frac{\partial \mathcal{L}}{\partial s_0}$, $\frac{\partial \mathcal{L}}{\partial \theta}$, $\frac{\partial \mathcal{L}}{\partial t_0}$ 和 $\frac{\partial \mathcal{L}}{\partial t_n}$ 。

2.1 共轭态方法

共轭态方法 [Plessix(2006), Chen *et al.*(2018)Chen, Rubanova, Bettencourt, and Duvenaud] 是专门针对积分过程反向传播的传统方法。在研究中, 人们发现积分过程的导数的反向传播同样是一个积分过程, 只不过方向相反。于是人们通过构造一个可以同时更新原函数和导数的拓展函数, 以对拓展函数的逆向积分的形式完成导数的计算, 如算法 1所示。

算法 1: 共轭态法

输入: 动力学参数 θ , 开始时间 t_0 , 结束时间 t_n , 末态 s_n , 以及需要回传的导数 $\frac{\partial \mathcal{L}}{\partial s_n}$
输出: $\frac{\partial \mathcal{L}}{\partial s_0}$, $\frac{\partial \mathcal{L}}{\partial \theta}$, $\frac{\partial \mathcal{L}}{\partial t_0}$, $\frac{\partial \mathcal{L}}{\partial t_n}$

```

1  $\frac{\partial \mathcal{L}}{\partial t_n} = \frac{\partial \mathcal{L}}{\partial s_n}^T f(s_n, t_n, \theta)$  # 计算损失函数对终了时间的导数
2 function aug_dynamics([s, a, -, -], t,  $\theta$ ) # 定义拓展动力学函数
3    $s' = f(s, t, \theta)$ 
4   return ( $s'$ ,  $-a^T \frac{\partial s'}{\partial s}$ ,  $-a^T \frac{\partial s'}{\partial \theta}$ ,  $-a^T \frac{\partial s'}{\partial t}$ )
5 end
6  $S_0 = (s_n, \frac{\partial \mathcal{L}}{\partial s_n}, 0, -\frac{\partial \mathcal{L}}{\partial t_n})$  # 计算拓展动力学函数的初始状态
7  $(s_0, \frac{\partial \mathcal{L}}{\partial s_0}, \frac{\partial \mathcal{L}}{\partial \theta}, \frac{\partial \mathcal{L}}{\partial t_0}) = \text{ODESolve}(S_0, \text{aug\_dynamics}, t_n, t_0, \theta)$  # 对拓展动力学反向积分
```

该算法的描述来自文献 [Chen *et al.*(2018)Chen, Rubanova, Bettencourt, and Duvenaud], 其中可以找到详细的推导过程, 这里对原算法中的符号做了替换以方便读者理解。

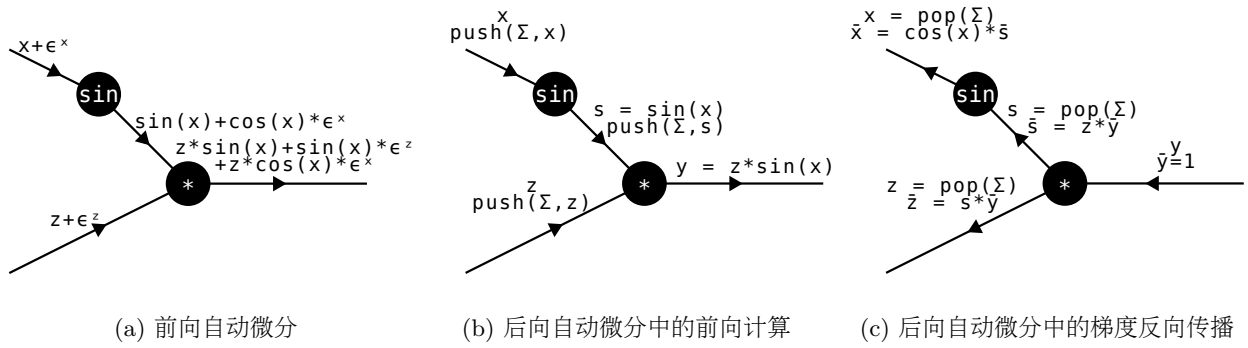


图 1: 利用自动微分计算对于计算过程 $y = z * \sin(x)$ 的导数 $\bar{x} \equiv \frac{\partial y}{\partial x}$ 和 $\bar{z} \equiv \frac{\partial y}{\partial z}$ 。其中圆圈代表函数，线代表变量，箭头代表运算方向， Σ 是一个用于缓存中间计算结果的全局堆栈（一种以先进后出的存储结构）而 **push** 和 **pop** 分别代表了对该堆栈的入栈（存储）和出栈（取出）操作。

2.2 前向自动微分

自动微分可以分为两大类，分别是前向传播 (Forward propagation) [Wengert(1964)] 和后向传播 (Backward propagation) [Boltysanski *et al.*(1960) Boltysanski, Gamkrelidze, Mishchenko, and Pontryagin]。

前向自动微分和数学中的链式法则推导导数的方法极为相似。它通过修改程序运算中的变量，使之携带若干无穷小量 ϵ^i ，并通过这个无穷小量的运算规则完成对程序的求导。当函数 f 作用于一个标量，它的运算规则变为

$$f(x + \sum_{i=1}^k y_i \epsilon^i) = f(x) + \frac{\partial f(x)}{\partial x} \sum_{i=1}^k y_i \epsilon^i \quad (2)$$

这里的上标 i 代表它对应第 i 个变量。容易验证这个无穷小量的运算规则与数学分析中的无穷小量并无差别。程序中仅记录一阶小量的系数，因此无穷小量满足运算规则 $\epsilon^i \epsilon^j = 0$ 。值得注意的是，前向自动微分的计算时间随着需求导的变量的数目线性增长。因为随着变量数目的增加， ϵ^i 的数目线性增加，函数的操作数也随之线性增加。图 1 (a) 所示的是对函数 $y = z * \sin(x)$ 的前向自动微分，这个函数仅仅包含两步运算 \sin 与 $*$ 。程序中每个变量都记录了三个域，分别是（变量本身， ϵ^x 的系数， ϵ^z 的系数），每当经过一个函数便更新数值和两个无穷小量的系数，因此整个函数的运行空间和一次求导的变量的个数也成正比。但在实践中，由于内存大小有限且需要求导变量较多的情况下，一般不会一次性的对所有变量都求导。而是重复运行多次计算过程，一次只处理若干个变量的导数这样分批求导。例子中这样简单的表达式对于人类来说处理起来也毫无难度，但真实的程序可能会包含数以亿计的这样的基础操作，虽然结果依然是解析的，但是人们很难通过人力得到解析的导数表达式。然而计算机恰恰很擅长这样繁琐但是规则简单的任

务。在前向自动微分框架中，单步的对控制参数 θ 的前向自动微分可表达为

$$(s_{i+1}, \frac{\partial s_{i+1}}{\partial \theta}) = (\text{ODEStep}(s_i, t_i, \theta, \Delta t), \frac{\partial s_{i+1}}{\partial s_i} \frac{\partial s_i}{\partial \theta})$$

其中第二个域为一阶无穷小量的系数（导数），由于状态 s_i, s_{i+1} 和控制参数 θ 均可包含多个变量，上述偏微分均解释为雅可比行列式，真实的计算中，并不需要构造出具体的局域雅可比矩阵 $\frac{\partial s_{i+1}}{\partial s_i}$ 并做复杂度较高的矩阵乘法，而是利用代码的变换或对基本运算符的重载来实现。

2.3 后向自动微分用同样的例子展示了后向自动微分的前向计算过程和梯度反向传播过程，前向计算过程中除了运算本身，还会将部分运算的中间结果通过 $\text{push}(\Sigma, \mathbf{x})$ 操作压栈将一个变量 \mathbf{x} 其放入一个全局堆栈 Σ 中，而后向过程将这些缓存的结果通过 $\mathbf{x} = \text{pop}(\Sigma)$ 操作出栈并利用如下递推公式计算导数

$$\bar{x} = \bar{y} \frac{\partial f(x)}{\partial x} \quad (3)$$

其中， $\bar{\cdot}$ 代表了损失对相应变量的微分 $\frac{\partial \mathcal{L}}{\partial \cdot}$ ，这里 \mathcal{L} 为标量损失函数，且 $\frac{\partial \mathcal{L}}{\partial \mathcal{L}} \equiv 1$ 。导数回传的计算复杂度与需要求导的变量数目无关，因此它在求导 $\sim 10^2$ 以上的输入变量的优化问题中对比前向自动微分拥有优势，这也是为什么所有的主流自动微分框架都采用梯度后向传播的模式。后向自动微分也有缺点，那就是向堆栈中存储数据带来了正比于计算步骤数 ($\mathcal{O}(T)$) 的额外空间开销以及频繁访问内存带来的性能下降的问题，这是后向自动微分框架设计的复杂性的来源。在求解常微分方程的自动微分过程中，其前向计算过程（左侧）和反向传播过程（右侧）分别为

$$s_{i+1} = \text{ODEStep}(s_i, t_i, \theta, \Delta t)$$

$$s_i = \text{pop}(\Sigma)$$

$$\text{push}(\Sigma, s_i)$$

$$\bar{s}_i = \bar{s}_{i+1} \frac{\partial s_{i+1}}{\partial s_i}$$

如何设计算法在保证对计算状态逆序的访问的前提下，减少计算中使用的堆栈 Σ 的大小，是我们接下来要讨论的一个至关重要的问题。一种节省空间方案叫做检查点方案，比起将每一步的计算状态保存到堆栈，有时候我们也可以通过“重计算”的方式来得到一部分中间状态。

	$s_{i+2} = \text{pop}(\Sigma)$
	$\overline{s_{i+2}} = \overline{s_{i+3}} \frac{\partial s_{i+3}}{\partial s_{i+2}}$
$\text{push}(\Sigma, s_i)$	$s_i = \text{read}(\Sigma, i)$
$s_{i+1} = \text{ODEStep}(s_i, t_i, \theta, \Delta t)$	$s_{i+1} = \text{ODEStep}(s_i, t_i, \theta, \Delta t) \# \text{ 重计算}$
$s_{i+2} = \text{ODEStep}(s_{i+1}, t_i, \theta, \Delta t)$	$\overline{s_{i+1}} = \overline{s_{i+2}} \frac{\partial s_{i+2}}{\partial s_{i+1}}$
$\text{push}(\Sigma, s_{i+2})$	$s_i = \text{pop}(\Sigma)$
	$\overline{s_i} = \overline{s_{i+1}} \frac{\partial s_{i+1}}{\partial s_i}$

这里，**read** 函数仅读取数据而不对数据出栈。在正向计算过程中，我们选择性的没有存储状态 s_{i+1} 。在反向计算过程中，由于状态 s_{i+1} 不存在，程序通过读取 s_i 的状态重新计算得到 s_{i+1} 。通过这种方式，堆栈的大小可以减少 1。如何最优的选择存储和释放状态成了检查点方案的关键，这个最优方案在 1992 年被 Griewank [Griewank(1992)] 提出，被成为 Treeverse 算法，该算法仅需对数多个存储空间，即表 1 所示的空间复杂度 $\mathcal{O}(S \log(T))$ ，其中 S 为单个状态的空间大小， T 为时间，或这里的步数。以及对数的额外时间开销，即时间复杂度 $\mathcal{O}(T \log(T))$ 。

还有一种可以保证变量方案叫做可逆编程，意思是让代码在结构上具有可逆性，比如在可逆编程语言的框架下书写代码。可逆的书写意味着赋值和清除变量这些常见的操作不再被允许，取而代之的是累加 ($+=$) 和累减 ($-=$)，从内存空“借”一段空内存 (\leftarrow) 和将一个已经清除的变量“归还”给内存 (\rightarrow) 等。可逆计算赋予用户更高的自由度去控制内存的分配与释放，以如下的程序计算过程（左侧）和它的反过程（右侧）为例

$s_{i+1} \leftarrow 0$	$s_{i+1} \leftarrow 0$
$s_{i+1} += \text{ODEStep}(s_i, t_i, \theta, \Delta t)$	$s_{i+1} += \text{ODEStep}(s_i, t_i, \theta, \Delta t) \# \text{ 计算 } s_{i+1}$
$s_{i+2} \leftarrow 0$	$s_{i+2} -= \text{ODEStep}(s_{i+1}, t_i, \theta, \Delta t)$
$s_{i+2} += \text{ODEStep}(s_{i+1}, t_i, \theta, \Delta t)$	$s_{i+2} \rightarrow 0$
$s_{i+1} -= \text{ODEStep}(s_i, t_i, \theta, \Delta t) \# \text{ 反计算}$	$s_{i+1} -= \text{ODEStep}(s_i, t_i, \theta, \Delta t)$
$s_{i+1} \rightarrow 0 \# \text{ 归还 } s_{i+1}$	$s_{i+1} \rightarrow 0$

这里我们先是从状态 s_i 经过两步计算得到了状态 s_{i+1} 和 s_{i+2} 。由于 s_{i+1} 在接下来的运算中不会被

用到，我们可以利用已有信息 s_i 将它反计算为 0，并将存储空间“归还”给系统。最优的可逆计算框架下的时间与空间的权衡一般认为是 Bennett 算法，其空间的额外开销也是对数，但是时间的额外开销变为了多项式，即表 1 所示的时间复杂度 $\mathcal{O}(T^{1+\epsilon})$ ，其中 $\epsilon > 0$ ，高于最优检查点方案。

3 后向自动微分的时间与空间的权衡

Treeverse 算法和 Bennett 算法分别是源代码后向自动微分（一种通过源代码变换实现的反向传播技术）和可逆计算中权衡时间与空间的核心算法。为了更好的理解这些算法以方便更好的对物理模拟过程微分，我们用鹅卵石游戏模型来进行说明。

3.1 鹅卵石游戏

鹅卵石游戏是一个定义在一维格子上的单人游戏，它最初被提出描述可逆计算中的时间与空间的权衡。游戏开始，玩家拥有一堆鹅卵石以及一个一维排布的 n 个格子，标记为 $0, 1, 2 \dots n$ ，并且在 0 号格子上有

鹅卵石游戏-可逆计算版本

放置规则：如果第 i 个格子上有鹅卵石，则可以从自己堆中取一个鹅卵石放置于第 $i+1$ 个格子中，
回收规则：仅当第 i 个格子上有鹅卵石，才可以把第 $i+1$ 个格子上的鹅卵石取下放入自己的堆中，
结束条件：第 n 个格子上有鹅卵石。

这里一个鹅卵石代表了一个单位的内存，而放置和取回鹅卵石的过程分别代表了计算和反计算，因此均需要一个步骤数，即可逆计算的一个单位的运算时间。可逆计算在释放内存时，要求其前置状态存在以保证反计算的可行，在这里对应回收规则中要求前一个格点中存在鹅卵石。游戏目标是从自己的堆中取尽可能最少的鹅卵石，或是使用尽可能少的步骤数触发游戏结束。可逆计算版本的鹅卵石游戏最省步骤数的玩法，也是最消耗空间的玩法是用 n 个鹅卵石依次铺至终点格子 n ，此时时间复杂度和空间复杂度均为 $\mathcal{O}(T)$ ，此处 T 为格子数。最少的鹅卵石数目的玩法则需要用到可逆计算框架下时间和空间最优交换方案 Bennett 算法。

如算法 2（图 5 (b)）所示，Bennett 算法将格子均匀的分割为 $k \geq 2$ 等份，先是像前执行 k 个区块得到计算结果，然后从最后第 $k-1$ 个区块开始依次收回中间 $k-1$ 个鹅卵石到自由堆中。每个区块又递归的均匀分割为 k 个子分块做同样的放置鹅卵石-保留最后的鹅卵石-取回鹅卵石的操作，直到程序无法再分

算法 2: Bennett 算法

输入: 初始状态集合 $S = \{0 : s_0\}$, 子分块数目 k , 分块起点 $i = 0$, 分块长度 $L = n$

输出: 末态 $S[n]$

```

1 function bennett( $S, k, i, L$ )
2   if  $L = 1$  then
3      $S[i + 1] \leftarrow 0$ 
4      $S[i + 1] += f_i(S[i])$ 
5   else
6     for  $j = 1, 2, \dots, k$  do
7        $\text{bennett}(S, k, i + \frac{j-1}{k}L, \frac{L}{k})$            # forward for k sub-blocks
8     end
9     for  $j = k - 1, k - 2, \dots, 1$  do
10       $\sim\text{bennett}(S, k, i + \frac{j-1}{k}L, \frac{L}{k})$          # backward for k-1 sub-blocks
11    end
12  end
13 end
  
```

割。假设次过程的递归次数为 l , 我们可以得到步骤数和鹅卵石与 k 和 l 的关系如下

$$T_r = (2k - 1)^l, S_r = l(k - 1). \quad (4)$$

其中, k 与 l 满足 $T = k^l$ 。可以看出可逆计算的时间复杂度和原时间为多项式关系。同时可以看出 k 越小, 使用的总鹅卵石数目越小, 因此最省空间的鹅卵石游戏解法对应 $k = 2$ 。作为例子, 图 3 (b) 展示了 $n = 16, k = 2$ ($l = 4$) 时候的游戏解法, 对应步骤数为 $(T_r + 1)/2 = 41$, 这里的实际操作数少了大约一半是因为最外层的 Bennett 过程不需要取回鹅卵石的过程。

我们稍微修改可以得到检查点版本的规则, 它为用户增加了一支画笔用于涂鸦格点, 改变后的规则描

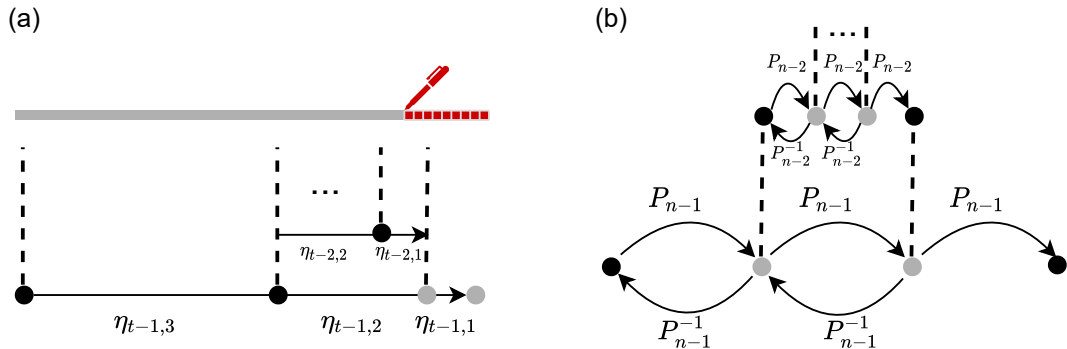


图 2: (a) 广义自动微分中常见的 Treeverse 算法 [Griewank(1992)], 其中 $\eta(\tau, \delta) \equiv \binom{\tau + \delta}{\delta} = \frac{(\tau + \delta)!}{\tau! \delta!}$ 。 (b) 可逆计算时空交换的 Bennett 算法。 [Bennett(1973), Levine and Sherman(1990)] 其中, P 和 Q 分别代表了计算和反计算。

述为

鹅卵石游戏-检查点版本

放置规则：如果第 i 个格子上有鹅卵石，则可以从自己堆中取一个鹅卵石放置于第 $i+1$ 个格子中，

回收规则：可以随意把格子上的鹅卵石取下放入自己的堆中，收回鹅卵石不计步骤数，

涂鸦规则：当第 i 个格子有鹅卵石，且第 $i+1$ 个格子被涂鸦或 $i=n$ ，可以涂鸦第 i 个格子，涂鸦不记入步骤数，

结束条件：涂鸦完所有的格点。

算法 3: Treeverse 算法

输入：初始状态 $s = s_0$ ，状态缓存集合 $S = \{\}$ ，需回传的梯度 $\overline{s}_n \equiv \frac{\partial \mathcal{L}}{\partial s_n}$ ，允许缓存的状态数 δ ，扫描次数 τ ，分块起点 $\beta = 0$ ，分开终点 $\phi = n$ ，以及把分块分割为两部分的分割点 $\sigma = 0$

输出：回传的梯度 $\overline{s}_0 \equiv \frac{\partial \mathcal{L}}{\partial s_0}$

```
1 function treeverse( $s, S, \overline{s}_\phi, \delta, \tau, \beta, \sigma, \phi$ )
2   if  $\sigma > \beta$  then
3      $\delta = \delta - 1$ 
4      $S[\beta] = s_\beta$  # 缓存状态  $s_\beta$  至状态集合  $S$ 
5     for  $j = \beta, \beta + 1, \dots, \sigma - 1$  do
6        $s_{j+1} = f_j(s_j)$  # 计算  $s_\sigma$ 
7     end
8   end
9   # 以  $\kappa$  为最优分割点（二项分布），递归调用 Treeverse 算法
10  while  $\tau > 0$  and  $\kappa = \lceil (\delta\sigma + \tau\phi) / (\tau + \delta) \rceil < \phi$  do
11     $\overline{s}_\kappa = \text{treeverse}(s_\sigma, S, \overline{s}_\phi, \delta, \tau, \sigma, \kappa, \phi)$ 
12     $\tau = \tau - 1$ 
13     $\phi = \kappa$ 
14  end
15  if  $\phi - \sigma \neq 1$  then
16    error("treeverse fails!") # 由于总步长  $n \neq \eta(\tau + \delta, \delta)$ ，实际实现应处理该异常
17  end
18   $\overline{s}_\sigma = \overline{s}_{\sigma+1} \frac{\partial s_{\sigma+1}}{\partial s_\sigma}$  # 利用已有的  $s_\sigma$  和  $\overline{s}_\phi$  回传导数
19  if  $\sigma > \beta$  then
20    remove( $S[\beta]$ ) # 从缓存的状态集合中移除  $s_\beta$ 
21  end
22  return  $\overline{s}_\sigma$ 
23 end
```

检查点版本的鹅卵石游戏中，涂鸦过程代表了梯度反向传播的过程，因此它要求按照与程序正常运行方向相反的顺序访问计算状态。它最节省步骤数的解法和可逆计算版本一样，即计算过程中不取下任何鹅卵石。而用最少鹅卵石的解法则是每当我们需涂鸦一个格子 i ，我们总是从初始鹅卵石 s_0 开始扫描

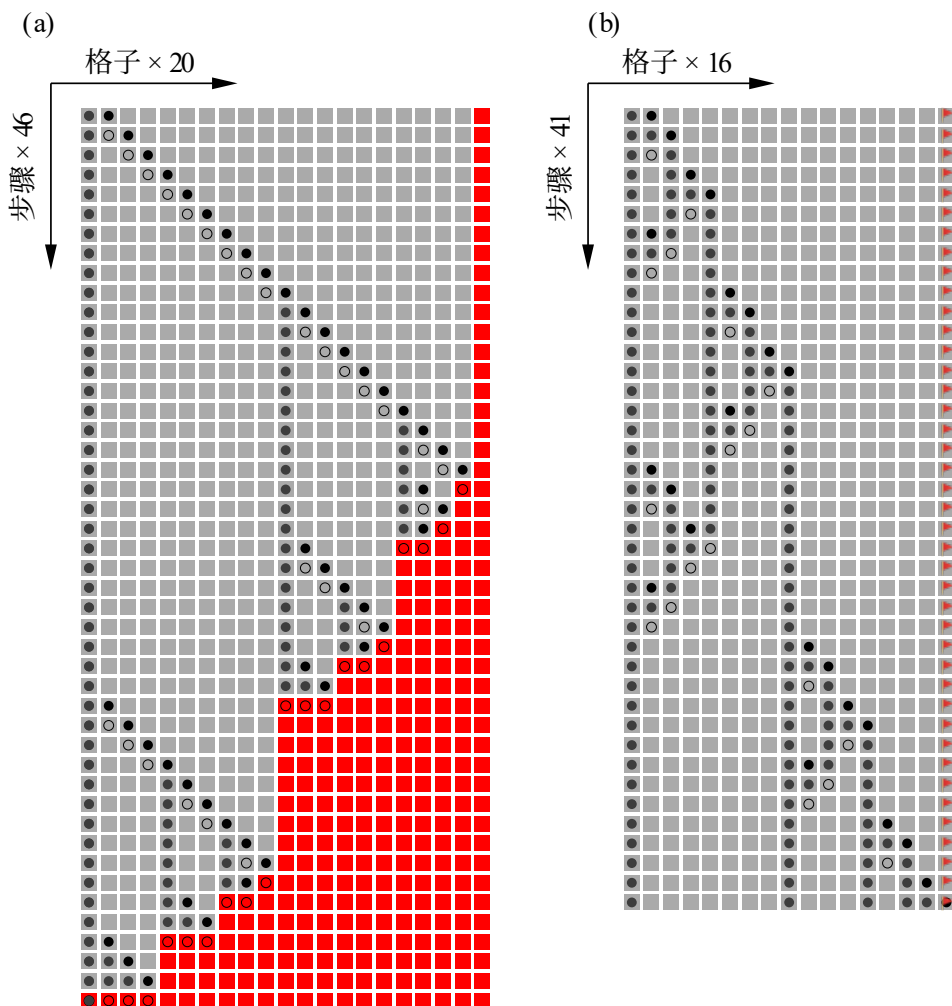


图 3: (a) Treeverse 算法 ($t = 3, d = 3$) 和 (b) Bennett 算法 ($k = 2, n = 4$) 对应的时间空间交换策略下的鹅卵石游戏，横向是一维棋盘的格子，纵向是步骤。其中“ \circ ”为在这一步中收回的鹅卵石，“ \bullet ”为在这一步中放上的鹅卵石，而颜色稍淡的“ \bullet ”则对应遗留在棋盘上未收回的鹅卵石。红色格子代表已被涂鸦，带旗帜的格点代表终点。

方法	时间	空间
共轭态法	$\mathcal{O}(T)$	$\mathcal{O}(TS)$
前向自动微分	$\mathcal{O}(NT)$	$\mathcal{O}(S)$
基于检查点的后向自动微分	$\mathcal{O}(T \log T)$	$\mathcal{O}(S \log T)$
基于可逆计算的后向自动微分	$\mathcal{O}(T^{1+\epsilon})$	$\mathcal{O}(S \log T)$

表 1: 不同方案的时间与空间复杂度。其中共轭态法考虑了缓存部分中间结果以保证反向积分的正确性, 因此空间会有与时间线性增长。前向自动微分中的 N 代表了需要求导的参数个数。可逆计算中的时间复杂度为多项式, 且 $\epsilon > 0$ 。

(依次放置一个鹅卵石并取下前一个鹅卵石) i 步至格子 i , 因此只需要 2 个鹅卵石即可涂鸦全部格子, 步骤数为 $\frac{n(n-1)}{2}$ 。如算法 3 (图 5 (a)) 所示, 完成第一遍从 s_0 到 s_n 的扫描后会在棋盘上留下 δ 个鹅卵石 (不包括初始鹅卵石), 把格点分割成 δ 个区块。我们把这些没有被取下的鹅卵石称为检查点, 我们总可以从任意一个检查点出发通过放置-取回鹅卵石的操作扫描后方的格子。每个区块的大小被证明有一个最优的取值, 即大小为二项分布函数 $\eta(\tau, d)$, 其中 $d = 1, 2, \dots, \delta$ 为从末尾开始数的区块的指标, 而 τ 的取值满足 $\eta(\tau, \delta) = n$ 。由于最后的 n 号格子可以直接被涂色, 拥有状态 s_{n-1} 后, $n-1$ 号格子满足涂鸦规则, 因此我们可以在第一遍扫描时给它涂上颜色。为了继续涂鸦 $n-2$ 号格点, 我们从离 $n-2$ 号格点最近的检查点出发扫描至该点, 依次类推直至达到最后一个检查点处。由于最后一个区块尺寸最小, 我们并不担心这样的扫描会使得步骤数增加太多。当我们完成了最后一个区块的涂鸦, 我们便可把格子上用于标记最后一个区块起点的鹅卵石取下重复利用。为了涂鸦倒数第二个区块, 我们先是扫描整个区间, 并把这个区间用回收的鹅卵石依据二项分布 $\eta(\tau-1, d=2, 1)$ 分割为两个子区间。随后依然是用同样的方式计算最后一个区间并递归的分割前一个子区间直至区块大小为 1 而无法继续分割, 同时也意味着进行了 τ 次递归。整个算法的时间和空间开销的关系是

$$T_c \approx \tau T, S_c = (\delta + 1)S, \quad (5)$$

其中, $T = \eta(\tau, \delta)$ 是初始计算时间。选择适当的 t 或者 d , 在时间和空间维度上的额外复杂度可以都是 $\log(T)$ 。图 3 (a) 展示了如何只用 4 个鹅卵石, 步骤数 46 涂鸦完所有 20 个格子。

图 4 展示了在固定额外时间开销的情况下, Bennett 算法和 Treeverse 算法得到的最优的空间开销。由于可逆性的限制, 可逆计算整体上需要更多的空间开销, 尤其是当步骤数更多, 或是允许的时间的额外开销更大的时候。但可逆计算也有优点, 其一是可以利用可逆性节省内存, 比如稀疏矩阵的运算中, 大多数运算都是可逆的基础操作。同时由于没有全局堆栈以及对程序自动设置检查点的问题, 程序设计上也更加自由, 比如在 GPU 编程的设备函数中是不允许访问全局堆栈的。基于鹅卵石模型的讨论对于通用的程序显然是过于理想化的, 但这样理想化的描述恰巧非常合适用于常微分方程微分的描述。

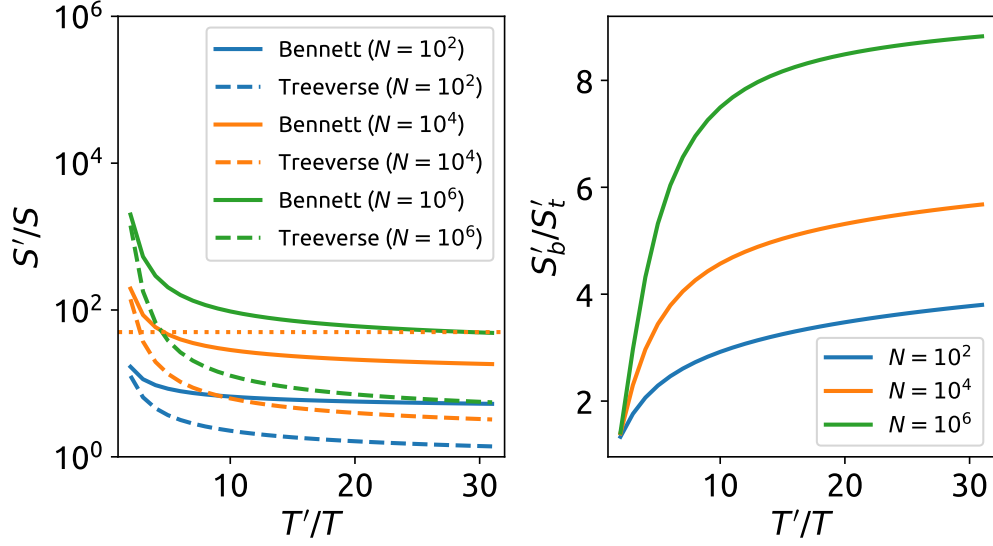


图 4: 为了回溯中间状态, 时间和空间在两种最优时间-空间交换策略下的关系。(a) 固定横轴为状态回溯的计算时间与原函数计算时间的比值, 对比再允许固定时间开销下, 内存的额外开销。其中 Bennett 算法代表了可逆计算下的最优策略, 而 Treeverse 则是传统计算允许的最优策略, 黄色点状横线对应 $S'/S = 50$ 。(b) 对比 Bennett 算法与 Treeverse 算法空间开销的比值。

4 自动微分在物理模拟中的应用

4.1 地震波的模拟

Perfectly matched layer (PML) 方程是模拟波在介质中运动的一种准确可靠的方案, 在介质中, 波场 $u(\vec{x}, t)$ 的传播可描述为

$$\begin{cases} u_{tt} - \nabla \cdot (c^2 \nabla u) = f & t > 0, \\ u = u_0 & t = 0, \\ u_t = v_0 & t = 0. \end{cases} \quad (6)$$

经过对空间的离散化处理后来在地震学的模拟中, 它被用来微分地震波传播的过程 [Symes(2007)]。自动微分也在其中发挥着重要的作用 [Zhu et al.(2020)Zhu, Xu, Darve, and Beroza]。[Grote and Sim(2010)]

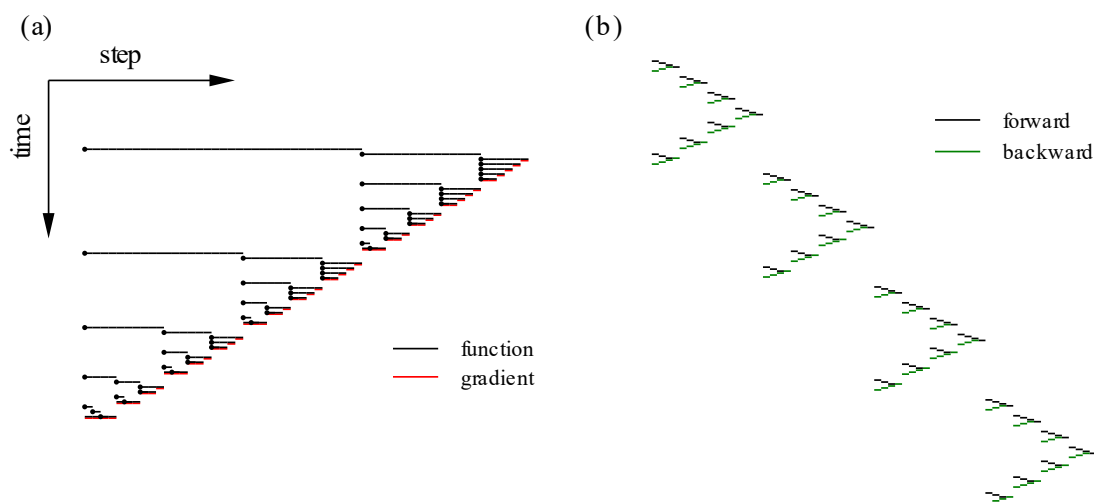


图 5: (a) Treeverse 算法 ($t = 5, d = 3$) 和 (b) Bennett 算法 ($k = 4, n = 3$) 中, 函数的执行过程, 其中横向是鹅卵石游戏的格子, 纵向随着进入 Treeverse 函数或者 Bennett 函数的次数向下延展。

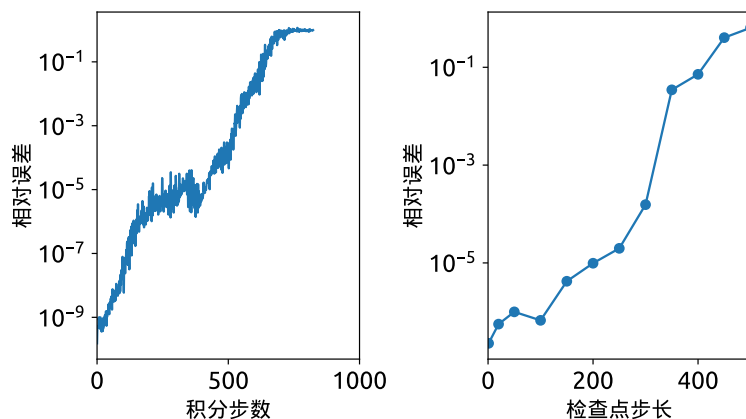


图 6: 利用共轭态方法求导时, l^2 误差与积分步长的关系。其中一个点代表了在该步长下, 对 100 个随机初始点计算得到的中位数。缺失的数据代表该处出现数值溢出的情况。

5 结 论

致谢感谢王磊老师的讨论。

附录 A2

[Rosset(2019)] C. Rosset, Microsoft Blog (2019).

[Nolan(1953)] J. F. Nolan, *Analytical differentiation on a digital computer*, Ph.D. thesis, Massachusetts Institute of Technology (1953).

[Gutzwiller(1963)] M. C. Gutzwiller, \bibfield journal \bibinfo journal Phys. Rev. Lett.\ \textbf{\bibinfo volume 10},\ \bibinfo pages 159 (\bibinfo year 1963).

[Carleo and Troyer(2017)] G. Carleo and M. Troyer, \bibfield journal \bibinfo journal Science\ \textbf{\bibinfo volume 355},\ \bibinfo pages 602—606 (\bibinfo year 2017).

[Deng *et al.*(2017)Deng, Li, and Das Sarma] D.-L. Deng, X. Li, and S. Das Sarma, \bibfield journal \bibinfo journal Physical Review X\ \textbf{\bibinfo volume 7} (\bibinfo year 2017),\ 10.1103/physrevx.7.021021.

[Cai and Liu(2018)] Z. Cai and J. Liu, \bibfield journal \bibinfo journal Phys. Rev. B\ \textbf{\bibinfo volume 97},\ \bibinfo pages 035116 (\bibinfo year 2018).

[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]

.

[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.

[Liu *et al.*(2020)Liu, Wang, and Zhang] J.-G. Liu, L. Wang, and P. Zhang, “Tropical tensor network for ground states of spin glasses,” (2020), arXiv:2008.06888 [cond-mat.stat-mech] .

- [Zhang *et al.*(2019)Zhang, Wan, and Yao] S.-X. Zhang, Z.-Q. Wan, and H. Yao, “Automatic differentiable monte carlo: Theory and application,” (2019), arXiv:1911.09117 [physics.comp-ph] .
- [Heimbach *et al.*(2005)Heimbach, Hill, and Giering] P. Heimbach, C. Hill, and R. Giering, Future Generation Computer Systems **21**, 1356 (2005).
- [Symes(2007)] W. W. Symes, \bibfield journal \bibinfo journal Geophysics\ \textbf{\bibinfo volume 72,\bibinfo pages SM213 (\bibinfo year 2007)}.
- [Zhu *et al.*(2020)Zhu, Xu, Darve, and Beroza] W. Zhu, K. Xu, E. Darve, and G. C. Beroza, “A general approach to seismic inversion with automatic differentiation,” (2020), arXiv:2003.06027 [physics.comp-ph] .
- [Plessix(2006)] R.-E. Plessix, \bibfield journal \bibinfo journal Geophysical Journal International\ \textbf{\bibinfo volume 167,\bibinfo pages 495 (\bibinfo year 2006)}, <https://academic.oup.com/gji/article-pdf/167/2/495/1492368/167-2-495.pdf> .
- [Chen *et al.*(2018)Chen, Rubanova, Bettencourt, and Duvenaud] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, in \emph{\bibinfo booktitle Advances in Neural Information Processing Systems, Vol. 31}, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., 2018).
- [Grote and Sim(2010)] M. J. Grote and I. Sim, “Efficient pml for the wave equation,” (2010), arXiv:1001.0319 [math.NA] .
- [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).
- [Griewank(1992)] A. Griewank, \bibfield journal \bibinfo journal Optimization Methods and software\ \textbf{\bibinfo volume 1,\bibinfo pages 35 (\bibinfo year 1992)}.
- [Bennett(1973)] C. H. Bennett (1973).

[Levine and Sherman(1990)] R. Y. Levine and A. T. Sherman, \bibfield journal \bibinfo journal SIAM
Journal on Computing\ \textbf{\bibinfo volume 19,\ \bibinfo pages 673 (\bibinfo year 1990)}.

Automatic differentiation in physics simulation *

Jin-Guo Liu¹⁾ Kai-Lai Xu²⁾

1) (Harvard University, Cambridge 02138)

2) (Stanford University, Stanford 94305)

1) (*Massachusetts Hall, Cambridge, MA 02138*)

2) (*450 Serra Mall, Stanford, CA 94305*)

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.

Keywords: automatic differentiation, scientific computing,
reversible programming, Treeverse, physics sim-
ulation

PACS: 02.60.Fh, 02.30.Jr, 91.30.-f

* 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: jinguoliu@g.harvard.edu