应用数学新时代的到来

鄂维南

1. Kepler 范式和 Newton 范式

自从 Newton (牛顿) 时代以来，做科学研究存在着两种不同的范式：Kepler (开普勒) 范式 (paradigm) 和 Newton 范式. 在 Kepler 范式或称为数据驱动的 (data driven) 方法中，人们通过对数据的分析提炼科学发现. 经典的例子是行星运动的 Kepler 定律. 生物信息提供了一个在当今时代引人注目的 Kepler 范式之成功的例证. 在 Newton 范式或称为基于第一性原理的 (first-principle-based) 方法中，目标是要发现支配我们周围的世界或我们所感兴趣的事物的基本原理. 最好的例子是理论物理，通过 Newton, Maxwell (麦克斯韦), Boltzmann (玻尔兹曼), Einstein (爱因斯坦), Heisenberg (海森伯) 和 Schrödinger (薛定谔) 的工作. 它现在仍然是一些最聪明大脑的一个重要用武之地.

数据驱动的方法随着统计方法和机器学习的发展已经成为一种强大的工具. 它在找到事实上是非常有效的，但是对于帮助我们找到事实背后的原因就没有那么有效了.

基于第一性原理的方法瞄准的是在非常基本的层面上的理解. 物理学，尤其是由对这种第一性原理的追求所驱动的. 1929 年量子力学的建立是一个拐点：如同 Dirac (狄拉克) 宣称 [2] 的那样，有了量子力学，对于大多数工程和特殊尺度物理学除外的自然科学，我们手上就已经有了必要的第一性原理.

但是，就像 Dirac 也指出的一样，描述量子力学的数学问题是极其复杂的. 困难之一是它是一个多体 (many-body) 问题：多加一个电子，问题的维度就增加了 3. 这是我们在基于第一性原理的方法中经常面对的困境：它是基本的，但不是非常实用的. 因此在实际中我们经常需要放弃严格和优雅的理论，转而诉诸临时而专门的非系统性的近似. 我们付出的代价不仅在于缺少了严格和优雅，还有结果的可信度和可移植性.

应用数学的发展也沿着相似的线. 由于物理学的第一性原理是用偏微分方程 (PDEs) 表述的，偏微分方程的分析和数值算法在应用数学中占据了一个中心的地位，特别是在上世纪 50 年代到 80 年代期间. 其目标是三重的：解决实际问题，理解它们背后的数学，为这些实际问题提供物理上的深刻见解. 一个非常吸引人的成功例子是流体力学. 流体力学不仅是偏微分方程研究的一个重要驱动力，流体力学研究在很大程度上已经成为一个计算的学科也是其发展出的数值算法的成功的证据. 对这些偏微分方程和数值算法的研究在应用数学中已经成为一个中心的主题很多年了，并且它今天仍然是一个活跃的领域.

当我还是加州大学洛杉矶分校的研究生时，我们被很骄傲地告知，我们属于“ Courant

译自：Notices of the AMS, Vol. 68 (2021), No. 4, p. 565–571, The Dawning of A New Era in Applied Mathematics, Weinan E.
Weinan E (鄂维南) 是美国普林斯顿大学的数学教授，他的邮箱地址是 weinan@math.princeton.edu.

(库朗) 式的应用数学 (Courant-style applied math)” 的阵营．创造这个词组是用来区别于“英国式的应用数学 (British-style applied math)” 的．两者都专注于流体力学．英国式的捍卫的是物理的洞察力和渐近理论．其领袖 Taylor (泰勒), Batchelor (巴彻勒), C. C. Lin (林家翘), Lighthill (莱特希尔) 等等，不仅是伟大的应用数学家，也是一流的理论流体力学家．大家也都知道他们一般不是特别重视数值计算和严格的分析．Courant 式捍卫的是数值计算和定理 (“定理证明者”)．其哲学是只要根本的偏微分方程和数值算法是可靠的，从计算中可以知道很多东西．毕竟，物理过程是非常复杂的；没有计算人们很难走得远．他们中的一些领袖，例如 von Neumann (冯 · 诺伊曼), Courant, Friedrich (弗里德里希斯), 以及 Lax (拉克斯), 不仅是伟大的应用数学家，也是伟大的纯粹数学家．这两个学派的长期不和被认为是应用数学中最大的敌意，这为流体力学在那个时代的主导地位做了辩护．

正统的数据驱动的研究团体是统计学界．不知为了什么原因，直到最近，统计学一直是非常独立于应用数学而发展的，并且事实上，独立于数学．一个数学系或者应用数学课程包含统计是非常罕见的．直至近些年才有了要求改变的呼声．

这不意味着应用数学界没对数据驱动的方法感兴趣．相反，自从上世纪 80 年代后期开始，随着小波和压缩感知 (compressed sensing) 的研究工作，信号和图形处理在应用数学中占据了一个中心的舞台．事实上，数据驱动方法的应用数学在过去 30 年间已经成为应用数学中最高产的领域之一．

这也不意味着流体力学是偏微分方程应用数学唯一成功的领域．事实上，一些人会说固体力学也是一样成功的：毕竟最重要的应用数学成功范例之一的有限元方法来自于固体力学．另一个成功的例子是数值线性代数：你只要看看 Matlab 有多流行就知道它的广泛影响了．这个名单还很长．

2．“Courant 式的应用数学” 的危机

不幸的是，在我这一代的 “ Courant 式的” 应用数学，流体力学的主导和成功，相对于机遇，提出的更多的是一种挑战．偏微分方程和流体力学的基础性工作已经由我们之前的几代人铺设好．我们剩下要做的要么是处理遗留的问题，例如湍流，要么是去征服新的领域．事实证明这二者都是困难的，更不用说复制应用数学在流体力学上的那种成功了．

实际上在流体力学之后， Courant 式的应用数学已经传播到很多其他的科学和工程学科中，例如材料科学，化学，生物学，神经科学，地学和金融，获得了很大的成功．但是一般来说，在这些领域中的成功度还达不到我们在流体力学中看到的水平．我们的贡献被欣然接受，但它们趋向于渐进式的，而非变革式的．因此，为处理他们面对的核心问题，科学家和从业者往往不得不求助于临时的既不可靠又不令人愉快的近似．这个情况在量子力学，分子动力学，粗粒化分子动力学 (coarse-grained molecular dynamics), 化学反应研究，复杂流体模型，可塑性模型中出现，蛋白质结构和动力学，湍流模拟，控制问题，动态规划等学科中也存在．

大多数 (如果不是全部) 这些问题的核心困难在于它们内禀上是高维问题，我们被维数灾难 (curse of dimensionality) 所困扰．

对于上面列出的大多数问题，高维度是问题的多尺度本质的结果，而多尺度，多物理学建模的想法带来了一丝希望．通过把小尺度上的不重要的自由度放在一起，人们应该能够直接使用更可靠的微尺度模型，为我们感兴趣的宏观尺度过程提出更高效的算法．尽管非常有希望，到目前为止，多尺度建模的成功是有限的，原因如下：

1. 微尺度模型它们本身往往不是那么可靠的．例如，在研究裂纹扩展 (crack propagation) 时，我们经常使用分子动力学作为微观模型．但是这些模型对于涉及断键过程的准确性通常是值得怀疑的．

2. 尽管多尺度建模可大大减少所需的微观模拟规模，但它仍超出我们目前的能力．

3．机器学习来拯救

上述困难事情的核心是我们处理多变量函数的有限能力，而这恰好是机器学习可以发挥作用的地方．通过提供近似多变量函数的能力，之前被认为不可能的事情现在变得极具可能性．

在机器学习之前，能常规化处理多变量函数的领域是数值积分．在统计物理中，我们几乎认为我们计算数百万个变量的函数积分的能力是理所当然的，而忘记了注意到这实际上是多么了不起．这是由多年来发展的蒙特卡洛算法 (Monte Carlo algorithm) 和方差减缩技术 (variance reduction techniques) 使之成为可能的．一点是，与基于网格的算法 (如 Simpson (辛普森) 法则) 不同，蒙特卡洛算法的收敛速率与维度无关．

逼近高维函数是一项更困难的任务, 机器学习的成功并不是容易的 [6]. 尽管神经网络模型很久以前就被发现了，直到最近才认识到它们在逼近多变量函数方面的全部潜力．然而，在很短的时间内，我们已经看到在机器学习的帮助下，在几个长期存在的问题上取得的一些显著的成就，还有很多有望在不久的将来出现．

将机器学习结合到应用数学中，将从根本上改变这两门学科．在下面，我们将讨论一些具体的例子来说明这将对科学计算，模拟，机器学习和纯数学产生的影响．

4．高维控制理论和偏微分方程

机器学习在科学计算中最早的成功应用之一是对高维控制问题的基于深度学习的 (deep-learning-based) 算法 [5]．首先，很有意思注意到一个问题，“维数灾难”一词是由 Richard Bellman (贝尔曼) 在动态规划的背景下首次提出的 [1]．事实上，Bellman 方程的维数与控制问题的状态空间的维数相同：如果我们对控制一个偏微分方程感兴趣，Bellman 方程就是无限维的．这严重限制了 “基于第一性原理的” 控制论的应用，许多实际问题不得不用特别的假设来解决，就像对量子多体问题一样．

在闭环控制的框架内，最优策略函数 (optimal policy function) 是状态的函数．如果用神经网络来参数化这个策略函数，那么随机控制和深度学习之间就有了一个非常好的相似性：控制问题的费用函数 (cost function) 是损失函数 (loss function); 控制问题的动力系统扮演着深度残差网络 (deep residual network) 的角色；动态系统中的噪声扮演着训练数据 (training data) 的角色，使我们能够使用随机梯度下降算法 (stochastic gradient descent algorithm) 进行训练. 有了这种基于深度学习的算法, 人们可以相当常规地处理数

百甚至更高维中的随机控制问题 [5]. 它也被扩展到确定性控制问题 (deterministic control problems)[7] 和一般非线性抛物型偏微分方程.

这些算法为处理现实世界中的控制问题和高维偏微分方程打开了大门. 这是一种令人兴奋的新的可能性，它将会影响 (并且在某种程度上已经影响) 经济学，金融学，运筹学和其他许多学科.

5. 机器学习辅助建模

在物理学中，我们习惯于基于第一性原理的模型. 这些模型不仅适用范围广，而且简单而优雅. Schrödinger 方程是一个很好的例子. 不幸的是，正如前面所指出的，使用这些模型解决实际问题可能是一项极其困难的任务. 由于这个原因，寻求简化模型一直是物理学和一般科学的一个永恒主题. 然而，正如我们在湍流模型中所经历的那样，如果我们不采用临时的特别的近似方法，通常很难得到这样的简化模型.

机器学习已经做好准备大力地提升我们发展这种基于物理的模型的能力. 这可以并且已经在发生，以 3 种不同的方式. 第 1, 它提供了能帮助多尺度模拟梦想成真的还缺失的工具. 第 2, 它提供了一个从数据直接发展模型的框架. 第 3, 顺着数据同化 (data assimilation) 的线，它可以提供一个将集成物理模型和观测相结合的强有力的工具.

然而，拟合数据是一回事，构建可解释的和真正可靠的物理模型则是非常不同的另一件事. 让我们首先讨论一下可解释性 (interpretability) 的问题. 众所周知，机器学习模型有一个 “黑盒子” 的名声，这为使用机器学习来帮助开发物理模型造成了心理障碍. 为了克服这个障碍，请注意，可解释性也应该从相对的角度来理解. 以气体动力学的 Euler 方程为例. 方程本身显然是可解释的，因为它们只代表质量、动量和能量的守恒. 但是，状态方程的细节是否可以被解释就不那么关键了. 事实上对于复杂气体，状态方程可能是以代码的形式出现的，它来自于使用样条插值的实验数据. 这些样条的系数是否可以被解释，并不是我们真正关心的问题. 同样的原则应该适用于基于机器学习的模型. 虽然这些模型的基本出发点应该是可解释的，就像气体动力学中的守恒定律一样，但进入这些模型的函数的具体形式并不都必须是可解释的. 这些函数通常代表一些构成关系，就像气体动力学中的状态方程.

现在转到可靠性问题上. 理想情况下，我们希望我们基于机器学习的模型能够像一般的物理模型 (如 Navier-Stokes (纳维 – 斯托克斯) 方程) 一样可靠，以达到所有的实际目的. 为了实现这一点，有两件事是至关重要的. 第 1 是基于机器学习的模型必须满足所有的物理约束，如来自于对称性和守恒定律的约束. 第 2 是我们用来训练模型的数据必须足够丰富，以充分代表实际遇到的所有物理情况. 由于标记数据几乎总是非常昂贵的，选择一个既小又具有代表性的好的数据集是开发这种模型的一个非常重要的组成部分. 我们将在下一节中详细讲这一点.

这些想法已经成功地应用于很多问题，包括分子动力学和稀薄气体动力学 [4]. 在分子动力学的情形，机器学习通过与高性能计算相结合，使得开始时的精度模拟具有数亿个原子的系统成为可能，提高了 5 个数量级 [4].

这些新的发展已经相当令人兴奋．但机器学习辅助建模的影响将在类似于生物学和经济学的领域中感受最深，在这些领域基于第一原理的建模很困难．在这些领域一些令人兴奋的进展正在发生着．

6. 机器学习的新前沿

机器学习与应用数学的结合也带来了一些机器学习的新机会．这里我们讨论两个．

6.1 并行 (concurrent) 机器学习 在大多数传统的机器学习情形中，训练数据要么是事先生成的，要么是被动观察得到的．当机器学习被应用于解决科学计算或计算科学中的问题时，情况通常不是这样的．在这些情形下，训练数据通常是即时生成的．与多尺度建模做一个类比，在多尺度建模中我们根据多尺度模型是事先生成的还是即时生成的来区分顺序 (sequential) 多尺度建模和并行 (concurrent) 多尺度建模，我们把这种机器学习的方式称为 *并行机器学习* (*concurrent machine learning*)．如同前面所看到的，生成一个最小但具有代表性的数据集是并行机器学习的一个关键问题．

为此，我们需要一个高效的步骤来探索状态空间，并需要一个标准来决定一个新遇到的状态是否应该被标记．一个例子是 [4] 中提出的 EELT 算法．

6.2 机器学习的 "适定" 表述 除了是令人惊讶地强大的之外，基于神经网络的机器学习也是相当脆弱的 —— 它的性能敏感地取决于模型中的超参数 (hyperparameters) 和训练算法．在许多情况下，参数的调整仍然是一门艺术，尽管随着经验的积累，这种情况正在稳步改善．

部分原因是，在机器学习中，模型和算法是在问题的表述被仔细思考之前构建的．就试想一下，如果我们不事先构建偏微分方程模型就试图对物理过程进行建模将会发生什么．事实上，一开始就有一个偏微分方程模型，并确保该偏微分方程模型是适定的，这是我们在 Courant 式的应用数学中学到的最重要的一课．

这激发了一个问题：我们能不能表述出机器学习的 "适定" 模型？期望是，如果我们从漂亮的连续表述开始，然后通过离散化来获得实用的模型和算法，那么对于超参数的选择，其性能将更加稳健．沿着这个思路，[3] 中已经做了一些初步的尝试．有趣的是，[3] 中工作的一个副产品神经网络模型是相当自然和不可避免的，因为最简单的连续模型和离散化得到的总是神经网络模型的这种或那种形式．尽管如此，这种处理机器学习的方式确实产生了新的模型和算法．更重要的是，它鼓励我们去寻找第一性原理，并允许我们在神经网络模型匣子之外思考．

在图像处理中可以找到一个很近的类比，就是去噪 (denoising)．标准的去噪方法是直接将精心设计的过滤器应用到图像上，看会发生什么．这种方法非常有效，特别是有了先进的基于小波的过滤器．另一种方法是写下一个去噪的数学模型，通常是以一个连续变分问题的形式，然后离散化，用优化算法解决离散化的模型．著名的 Mumford-Shah (芒福德–沙赫) 和 Rudin (鲁丁)-Osher-Fatemi 模型就是这种数学模型的例子．人们可能会质疑这些数学模型的正确性，但有一个定义明确的数学模型来开始显然有其优势．一方面，它帮助将图像处理变成有趣的偏微分方程问题．它还促使人们去思考图像处理背

后的基本原理，尽管沿着这条线没有取得多少进展.

希望是，这种新的数学理解和表述不仅有助于促进机器学习目前的成功，并且还能将其成功扩展到其他广泛的学科. 毕竟，机器学习是关于函数近似的，数学中一个非常基本的问题. 拥有在高维尤其有效的表示和逼近函数的新方法，将肯定会有重大而广泛的影响.

7. 高维分析

不仅是应用领域会受到影响，数学本身也会感受到影响，特别是分析.

机器学习带来了一系列高维中新的分析问题，从逼近函数到逼近概率分布，动力系统，以及求解偏微分方程和类似 Bellman 的方程. 研究这些问题将不可避免地产生一个数学中的新学科：高维分析.

在这个方向上，已经在数学中受到重视的一个领域是高维积分. 对蒙特卡洛方法的分析，特别是 Markov (马尔可夫) 链蒙特卡洛方法，在相当长的时间里一直是概率论和数学物理的一个活跃的领域.

积分是分析中最基本的问题. 人们可以就函数，概率分布，动力系统，变分法和偏微分方程问很多更加高级的问题. 例如，一个重要的问题是刻画这些对象的复杂度. 在一个抽象的层面上，复杂度应该由给定对象被简单的基本对象所近似的难度来定义. 例如，对于函数，基本对象可以是多项式、分片多项式或神经网络. 对于概率分布，基本对象可以是 Gauss (高斯) 分布的混合.

以函数的复杂性为例. 经典地，这是通过光滑度 (即函数可以被微分多少次) 做到的. 许多函数空间的等级都是按照这一思路定义的，比如 C^k 空间，Sobolev (索伯列夫) 空间和 Besov (别索夫) 空间. 在低维情形，这是很有道理的. 事实上，人们可以证明，这些空间中的函数是以当它们由某类基本函数 (如分片多项式或三角函数) 逼近时的收敛速度为特征的.

这种类型的结果是有维数灾难的. 事实上，越来越清楚的是，基于光滑度的概念并不是衡量高维函数复杂性的正确道路. 而我们应该通过是否能被一特定的类似神经网络的模型来高效逼近来衡量高维函数的复杂性. 通过这种方式，人们得到了再生核 Hilbert (希尔伯特) 空间 (reproducing kernel Hilbert space, RKHS), Barron 空间，多层空间和流诱导 (flow-induced) 空间，其中每一种都与一类特定的机器学习模型自然相关.

那么高维偏微分方程呢？一个自然的问题是，我们是否可以为上述函数空间中的某些类别的偏微分方程发展出一个正则性理论. 如果我们能做到，这意味着人们应该能够使用相应的机器学习模型来高效地逼近这些偏微分方程的解. 这个问题对于 Hamilton (哈密顿)–Jacobi (雅各比)–Bellman 方程尤其重要.

8. 应用数学作为一门成熟的学科

应用数学能否成为一个只有少数几个主要组成部分的统一学科，就像纯数学那样？我们能否有一个合理的统一的课程来培养应用数学家？这些问题长期以来一直难以解决. 回顾过去，很明显，当时的情况就是不成熟. 一方面，应用数学的确是非常多样化的，几

乎涉及到科学和工程的每一个学科. 寻求统一性和统一的课程无疑是一项艰巨的任务. 另一方面, 例如机器学习这样的主要组成部分在应用数学核心中的缺失的事实, 意味着它还没有准备好. 只要试想一下, 如果没有代数, 纯数学将会是什么样子.

情况已经改变. 随着机器学习出现在图景中, 应用数学的所有主要组成部分现在都已到位. 这意味着, 应用数学终于准备好成为一门成熟的科学学科. 是的, 新的方向将继续出现, 但有理由相信, 基本要素将不会有大的改变. 这些基本要素是: (基于第一性原理的) 建模、学习和算法.

8.1 应用数学的主要组成部分 代数, 分析, 几何和拓扑构成了纯数学的主要组成部分. 对于物理学来说, 它们是经典力学, 统计力学, 电磁学和量子力学. 应用数学的主要组成部分是什么? 以下是一个提议. 它并不一定是这个问题的最终说法, 而是进一步讨论的一个出发点.

应用数学有 3 个主要组成部分.

1. 基于第一性原理的建模, 它包括 (物理) 模型本身和对这些模型的分析工具. 简单地说, 前者是关于物理学, 后者是关于微分方程的.

物理模型背后的原理是物理学的基本定律和原理: 物理设定 (例如, 经典的还是量子的, 惯性主导的还是过度阻尼的), 变分原理, 守恒律, 等等.

这些第一性原理是以变分问题或微分方程的形式表述的. 因此, 我们需要分析工具来处理这些数学问题. 渐近方法可以迅速抓住问题的本质, 并给予我们急需的洞察力. 严格的定理可以帮助我们把事情建立在一个坚实的基础上, 此外还可以进一步说清楚该问题.

2. 数据驱动的方法. 到目前为止, 数据驱动的方法中最重要的部分是机器学习. 但也包括统计和数据 (如图像) 处理.

3. 算法. 在这里, 我们脑子中既有关于基于第一性原理的应用算法, 又有关于数据驱动的应用算法. 幸运的是, 这两个领域的算法有很多共同点. 一个例子是优化算法. 它们不仅在机器学习的成功背后发挥了举足轻重的作用, 许多基于第一性原理的模型被表述为需要优化算法的变分问题.

8.2 课程和教育 大多数 (如果不是全部) 一流大学都有相当成熟的纯数学本科和研究生课程. 很少有成熟的应用数学课程. 比这更糟糕的是, 在某些情况下, 应用数学课程被当作一套技巧来教, 而不是一个统一的科目. 一个例子是 “数学建模” 课程. 虽然这本应是一门应用数学的基础入门课程, 但它往往被当作一套例子来教, 没有一个一致的整体图像. 成熟的应用数学本科课程的缺乏是应用数学的一个最重要的障碍, 因为它阻碍了我们吸引年轻人才的能力.

明确了应用数学的主要组成部分, 我们现在可以为应用数学设计一个统一的课程. 自然地, 这个课程以上面讨论的 3 个主要组成部分为中心. 我们简要地讨论一下每个部分.

建模有两个部分: 模型的物理原理, 以及分析这些模型的数学工具. 前者如教给数

学家的物理学基础知识．后者是应用分析，包括常微分方程和偏微分方程，变分法，概率分布分析，渐近分析，和随机分析．每一个都可以由一年长的课程来覆盖．

学习真正的意思是数据分析．它由机器学习，数据处理和统计学组成．已经有适合应用数学家的数据处理和统计学的成熟课程．而机器学习的情况则不同．它例行地是以适合计算机科学家的方式进行教学的．我们需要一种方法将其教给数学家．在这一点上，机器学习在数学的方面还不是一个成熟的学科，但这种情况正在迅速改善．我们相信，一个合理的机器学习的数学介绍课程将很快被开发出来，它可以作为一个一学期的课程来教授．

算法有两个部分：对连续对象的算法和对离散对象的算法．前者由数学系提供的数值分析课程涵盖．后者由算法 / 离散数学课程涵盖，通常在计算机科学系教授．有了机器学习的发展，这两部分正在走到一起，所以以一种更统一的方式来教授它们是非常重要的．

开发所有这些课程将需要大量的努力，但我们应该并且能够使之实现．

9．应用数学作为跨学科研究的基础

有了这样的课程，应用数学将成为跨学科研究的基础．毕竟，建模，学习和算法是所有理论跨学科研究的基本组成部分．上述应用数学课程将有助于系统化对学生的训练和跨学科研究项目的组织．如果这成为现实，这将是跨学科研究历史上的一个转折点．

所有这些都需要时间．首先，我们需要从基础的事情开始，对年轻学生的培训．但在我们有机会培训他们之前，我们必须能够吸引他们进入应用数学．我在普林斯顿当了 20 多年的教员，对一件事印象非常深刻，就是数论如何能够吸引人才．我现在相信，令我自己都惊讶，应用数学也有潜力做到这一点．应用数学具有所有对年轻学生特别有吸引力的主要特征：问题的简单性和优雅性 (如机器学习) 以及这些问题带来的挑战 (如湍流), 还有一个额外的好处是，它是通向最令人兴奋的科学和技术的新发展的主要途径之一．我们在研究生层次开始看到这种变化．

在科学史上，有两个时期对应用数学的影响最大．第 1 个是 Newton 时代，在此期间，人们确定数学应该是科学的语言．第 2 个是 von Neumann 时期，在此期间，人们提出数值算法应该是数学和科学之间的主要桥梁．现在，第 3 个时期即将到来，在这个时代，应用数学的所有主要组成部分都已到位，形成跨学科科学研究的基础和令人激动的技术创新的基础．这真正是一个令人激动的时代．让我们一起努力，使之成为现实！

致谢 (略)

参考文献

[1] Richard Bellman, Dynamic programming, Princeton University Press, Princeton, N. J., 1957. MR0090477

[2] Paul A. Dirac, Quantum mechanics of many-electron systems, Proc. Roy. Soc. London Ser. A 123 (1929), no. 792.

[3] Weinan E, Chao Ma, and Lei Wu, Machine learning from a continuous viewpoint, I, Sci. China Math. 63 (2020), no. 11, 2233–2266, DOI 10.1007/s11425-020-1773-8. MR4170870

(下转 340 页)