

中国科学技术大学

“科学与社会”新生研讨课研究报告

报告题目： 使用神经网络模拟物理公式

小组组长： 修中淇

小组成员： 王俊棋 杨崇何

导师姓名： 邹旭波

2022 年 6 月 27 日

一、研究小组成员及其承担的主要工作

学号	姓名	所在学院	在研究和报告撰写中承担的主要工作
PB2102066 6	修中淇	物理学院	运行程序, 处理数据
PB2102066 7	王俊棋	物理学院	修正与勘误方法, 数据
PB2102066 4	杨崇何	物理学院	查找, 处理相关文献

二、进度安排

研究工作的具体时间节点和工作进度

上半学期: 2021 年 9 月——接收任务, 阅读论文, 开始相关能力学习

2021 年 12 月——修正代码, 测验数据运行成功

下半学期: 2022 年 3 月——接收进阶任务, 寻找黑体辐射数据

2022 年 6 月——最终运行得到理想结果

三、报告摘要

研究目的: 通过学习与运行应用程序最终模拟出黑体辐射的相关图像与理论公式。

方法: 程序使用了一种神经网络架构, 该架构可以模拟物理推理过程, 并可以用一种无偏的方式从实验数据中提取简单的物理概念。大多数物理规律都可以表述为偏微分方程的形式。如果神经网络具有至少一层非线性隐藏层, 只要网络有足够数量的神经元, 那么就可以充分地逼近任意一个在 R^n 的紧子集上定义的连续函数。

结果: 通过神经网络长时间的模拟, 在近似程度上得到了黑体辐射在不同温度下的图像, 因为模拟相对于实际有偏差, 得到的图像在一定范围内会有波动。由于阅读程序的能力有限, 并没有构建出较为迅速的模拟手段, 导致程序运行时间较长。

结论: 编译程序后借助 matlab 的图像显示功能, 可以通过神经网络模拟出黑体辐射的图像。

四、研究报告

用神经网络发现物理概念

在分析物理实验所得数据的过程中，研究者会受到当前发展时期流行的思想流派的影响，使用更符合当前理论的理论来解释实验现象，这样的分析方式在历史上获得了无数次的成功。但是，我们也许可以抛弃原有的物理基础知识，不强加任何已有的物理概念，来使用神经网络分析处理数据，发现经典和量子力学中的物理概念。通过几个阶段的共同学习与实践，我们的科学与社会研讨小组围绕科研论文〈Discovering physical concepts with neural networks〉，进行了研究和探索。

了解神经网络：生物的神经元分为

①输入（感受）区：传入信息的输入区，该区可以产生局部电位。

②整合（触发冲动）区：始段属于整合区或触发冲动区，众多的局部电位在此发生总和，并且当达到阈电位时在此首先产生动作电位。

③冲动传导区：轴突属于传导冲动区，动作电位以不衰减的方式传向所支配的靶器官。

④输出（分泌）区：轴突末梢的突触小体则是信息输出区，神经递质在此通过胞吐方式加以释放。

人工神经网络（Artificial Neural Network, ANN），正是基于生物学中神经网络的基本工作原理，在抽象了人脑结构和刺激响应机制后，模拟人脑的神

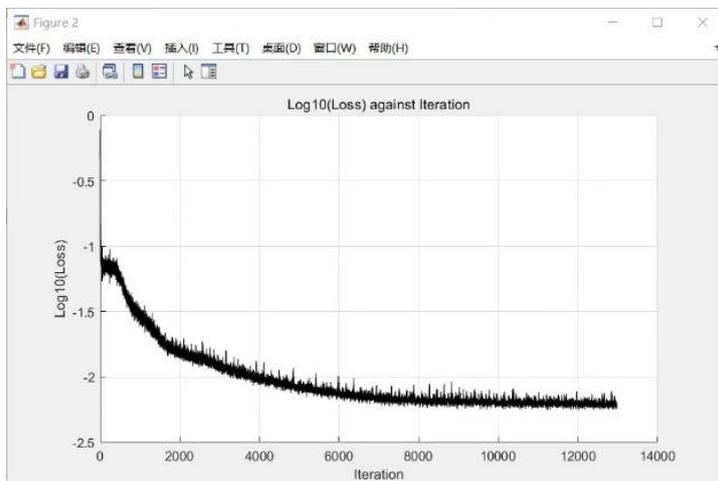
经系统来处理复杂信息的一种数学模型。在此模型中，各部分并行分布，具有高容错、自适应、智能化和自学习等特点。作为一种由大量简单元件相互连接而成的复杂网络，ANN 具有高度的非线性，从而能够实现非线性关系，进行复杂的逻辑操作，显然这正是机器学习研究的重要方向之一，也契合了我们探究物理规律，发现物理概念的需求。

此次实验所用的神经网络包含三层，输入层、隐含层、输出层（Input Layer/Hidden Layer/Output Layer），实验过程中，输入层编码器将实验所得的现象编码为实参数（real parameters），再将所求解的问题也作处理，编码为参数表示，随后共同输出到解码器中这些参数和表示被解码，输出一个结果。这种处理方式，与物理学中，人对数据的处理与分析类似，都是对处理过的数据和关系进行分析拟合，通过猜想函数关系来对结果进行预测。

论文展示了一种神经网络体系结构，简称为 SciNet，它可以模拟物理学家的建模过程，并将其应用于研究各种物理场景。例如简单的匀速运动，提供几组时间和位置，就可以得到任意时刻的位置。这个结构实质上是一种只能观测到输入和输出数据的黑盒（black box）分析结构，我们无法看到 SciNet 对数据和参数表示所作的处理，为了能够得到期望的物理规律，需要使用多组样本数据对此神经网络体系结构进行训练，原始条件被分类为现象（observations）、问题（questions）和答案（answers），分别编码（encode）为实参数，即 $(o, q, a_{cor}(o, q))$ 。SciNet 接收输入的实参数，对其进行加权计算，计算中，各参数的权值（weight）和最后数据的偏离值（biases）作为一组新的参数，在训练 SciNet 的过程中，允许它对这一组参数进行调整，以提高预测精准度，来拟合针对问题（Q）的答案

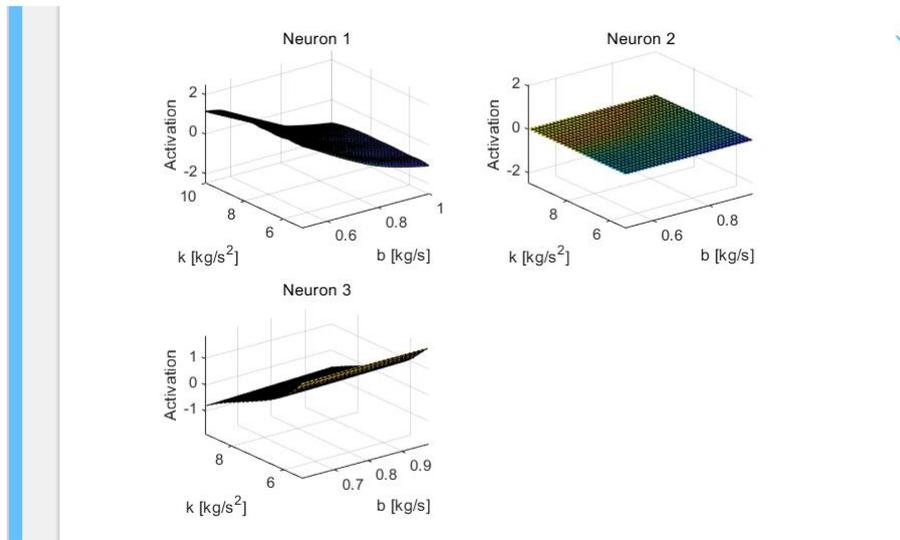
(A)。在此过程中，人工神经网络的复杂性和非线性性发挥优势，能够对数据做出全面的分析，找出其中内在的物理规律，发现守恒量和新的物理概念。为了使拟合结果更加准确，SciNet 的隐含层需要足够数量的神经元来模拟物理量之间的计算关系，且会定义一个。在经过多次自主学习和调整之后，只要此神经网络能够根据一些观测数据，对特定问题给出合理的答案，我们就可以通过它模拟物理过程，得到所需的物理信息。

通过对程序原理的初步了解，我们决定复现论文中较简单的一个实验，阻尼摆，通过在隐藏层建立数个神经元，然后向输入层输入多组时间(t)与位置(x)进行训练，SciNet 对数据进行了试探性的运算，最终确定位置变量 x 关于时间变量 t 需要三个参数，分别为 x 本身、x 对 t 的一阶导数的系数和常数项，同时，它也确定了参数具体的数值，这与牛顿力学中阻尼摆位移与时间的函数关系式 $kx+bx'+mx''=0$ 相吻合，可以证明我们对论文的理解基本正确。

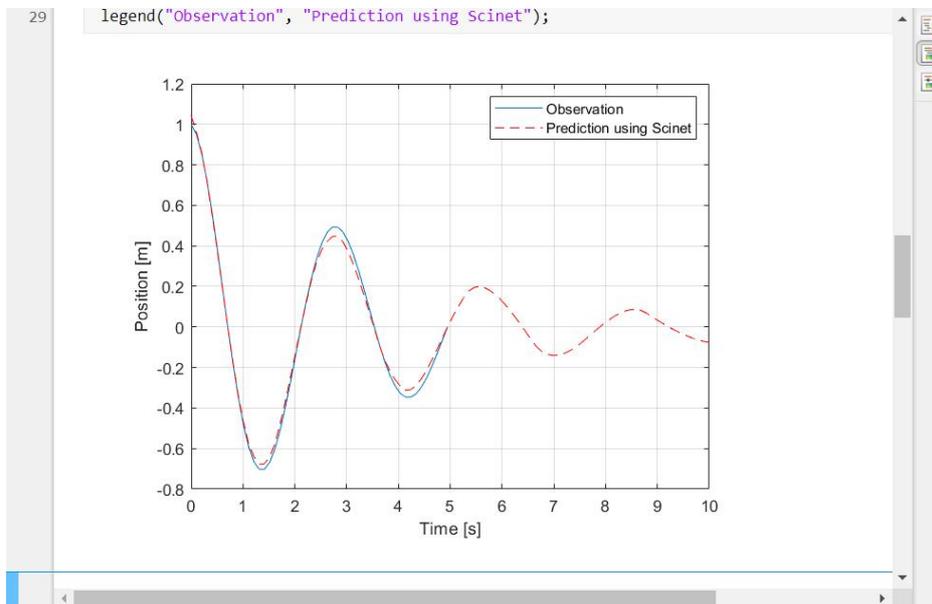


如图，程序运行过程中，随着迭代，程序以随机梯度下降法训练网络，使得潜在变量解缠，神经网络逐渐获得重构输入（observation）的能力，代价函数数值逐步下降。

程序运行后得到结果如下：



上图含义为潜在变量解缠后最终得到的三个相互独立的神经元，可见程序最终得到的潜在神经元数量为 3，其中两个神经元直接受变量影响，最后一个表示为单独的常量。



上图为神经网络拟合的预测图像，可以看出，拟合图像很好的反映了输入数据的性质和变化趋势，程序测验成功。

就神经网络的本质而言，我们学习到了以下两个方面：

1. 单一神经元

为实现人工神经元，我们使用了指数线性单位 (ELU) 为单一神经元的基本结构，它在输入变量的不同区间具有指数性质或线性性质且二者连接平滑，故神经网络的学习对象没有仅线性的局限。

神经网络是包含误差参数 θ 的，这意味着我们对数据处理时是精确的，而每次处理后程序会进行代表神经网络还原输入的程度的代价函数 (cost function) 的计算，并计算其对于 θ 的梯度值，从而得到下一次迭代 θ 的更新方向，这种算法称为随机梯度下降法。

2. 变分自动编码器 (VAEs)

SciNet 的实现使用了所谓的变分自动编码器 (VAEs) 的改进版本，标准的 VAE 体系结构不包括 SciNet 所使用的问题输入，并试图从反馈中重构输入。

具体来说，自动编码器使用一个神经网络 (编码器) 将输入 x 映射到少量的潜在神经元 z ，然后使用另一个神经网络 (译码器) 重建输入。在训练过程中，对编码器和解码器进行优化，使重构精度最大化。

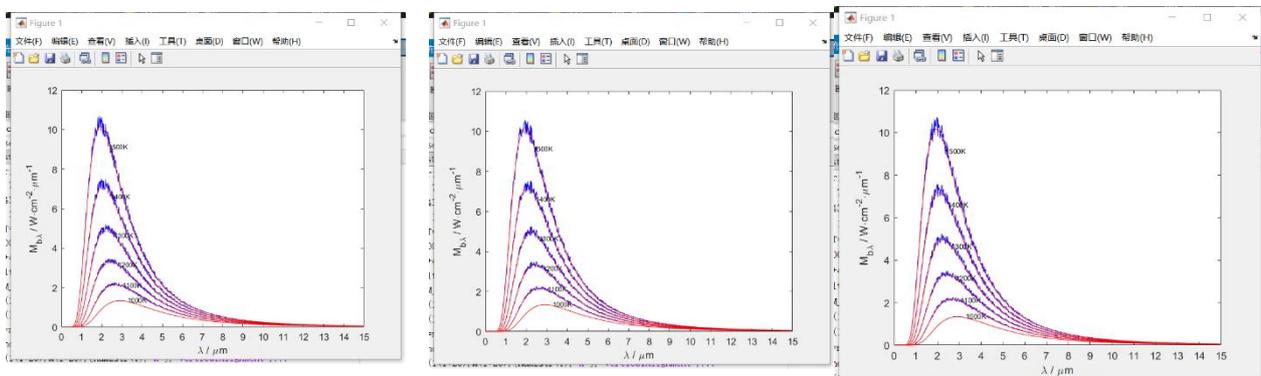
这个原理是通过先假设编码器分布为一种概率分布 (通常假设为独立的正态分布)，潜在向量从此分布中取样后补上权值和偏差，将结果映射给译码器 (为结果到输入的概率分布)，由其修正编码器和译码器的误差。

我们尝试将这种人工神经网络用于黑体辐射现象的研究。

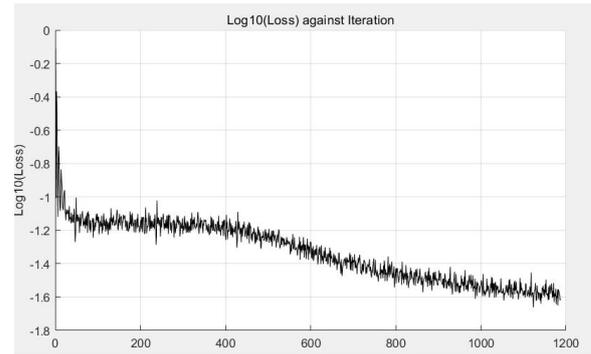
黑体辐射是物理学中一个很重要的问题。普朗克就是在研究黑体辐射问题的过程中提出量子假说的，从而掀起了物理学发展的又一个高潮。前辈们在这方面已经做出了许多杰出贡献。量子力学中提到的带小孔的空腔是一种理想黑

体。物理学家研究的是空腔中的电磁波辐射能量密度 U ，但是空腔中不仅充满了电磁场还包含着大量的空气分子，所以这里所说的黑体是指空气分子与电磁场处于热力学平衡的体系。我们在这里讨论的黑体就是这种热力学系统。如果加热腔壁，腔内壁就会发射电磁波，分子吸收电磁波后动能增加，气体温度升高，当分子吸收的电磁波和发射的电磁波相等的时候，两者就建立了平衡。这就是黑体升温的过程。现在我们再来看黑体辐射降温的过程。黑体向外辐射的时候，电磁波能量减小，黑体内部电磁场与空气分子之间出现了温差。这样就会使分子将储存的能量以电磁波的形式不断释放出来，以达到平衡。如果没有热源，黑体不断辐射就会使分子能量不断减小，这样黑体的温度就会就会降低。

实验：作为输入，我们提供了在特定温度下，不同波长时的黑体的光谱辐射出射度，并要求 SciNet 预测其他波长时的黑体的光谱辐射出射度。在没有给予其他物理概念的情况下，SciNet 会从模拟的其他波长序列数据中提取出黑体的光谱辐射出射度的相关物理参数。变换温度后，再次输入相应的波长与黑体的光谱辐射出射度，可以串行地在一个网络下得到不同温度下的拟合图像。如下



图所示。（图示为源码编译后的 MATLAB 程序运行结果）



我们可以看到，所画得的 figure 在两侧实际数据点的分布误差较大，在中间实际数据点分布较为稀疏的地方较为平滑。多次对相同程序进行运行，发现基本

的图像趋势较为完整，但每次模拟出的图像噪音不同，振动的幅度与位置也不同。根据误差图像的变化趋势来看，所得的模型逐渐接近于真实模型。

通过神经元的拟合结果，我们可以看到，在仅有波长 λ 一个变量时，M 满足的正比关系为 $M \propto \frac{1}{\lambda^5 (e^{\lambda} - 1)}$ 。当变化变量为 T 时，M 满足的正比关系为 $M \propto \frac{1}{e^{-aT} - 1}$ 。

之后我们根据拟合图像对相关系数进行分析，得到在所查阅数据的实验条件下正比关系的各项系数：

$$M(\lambda) = \frac{k_1}{\lambda^5 (e^{-k\lambda} - 1)} \text{ 其中 } k_1 \approx 3.773 \times 10^{-16}, k \approx 1.103 \times 10^{-5} (T = 1300K)$$

$$M(T) = \frac{k_2}{e^{-aT} - 1} \text{ 其中 } k_2 \approx 3.762 \times 10^{14}, a \approx 6.94 \times 10^{-5} (\lambda = 1\mu\text{m})$$

综合上式，我们可以得到

$$M(\lambda, T) = \frac{k'_1}{\lambda^5 (e^{k'_2 T \lambda} - 1)} \text{ 其中 } k'_1 \approx 3.773 \times 10^{-16}, k'_2 \approx 1.45 \times 10^{-2}$$

参考文献：

[1]Discovering physical concepts with neural networks

Raban Iten, Tony Metger, Henrik Wilming, L'idia del Rio, and Renato Renner

[2] Huembeli, P. , et al. "Automated discovery of characteristic features of phase transitions in many-body localization." (2018).

[3]Melnikov, et al. "Active learning machine learns to create new quantum experiments." Proceedings of the National Academy of Sciences of the United States of America (2018).

[4]周志华, and 陈世福. "神经网络集成." 计算机学报 25.1(2002):8.

[5] Ruotsalo, T. , et al. "SciNet." the 38th International ACM SIGIR Conference ACM, 2015.

学习体会：

经过两学期的程序理解与查找文献，模拟图像，在忙碌之余也得到了颇多的收获。我们学习了 Python 语言，学习了 MATLAB 的使用方法，并深深体会到机器学习与神经网络在理论物理方面给我们的学习带来了很大的方便。我认为，在这个研究中，在收获知识的同时，还收获了阅历，收获了成熟，在此过程中，我们通过查找大量资料，请教老师，以及不懈的努力，不仅培养了独立思考、动手操作的能力，在各种其它能力上也都有了提高。更重要的是，我们学会了很多学习的方法。而这是日后最实用的，真的是受益匪浅。要面对社会的挑战，只有不断的学习、实践，再学习、再实践。

五、导师评审意见

导师根据报告的整体质量、平时讨论时的表现、答辩的情况及在报告中各学生的贡献程度，着重考察学生的科学探索精神、自主学习能力、独立思考能力、

逻辑推理能力、实际动手能力、团队合作精神等要素，给小组成员做出评价。

导师签字：

日期：