基于奇异值分解的矩阵手术

Jehan Ghafuri 和**Sabah Jassim**白金汉大学
计算机学院，英国白金汉，MK18 1EG
{1901699,
sabah.jassim}@buckingham.ac.uk

arXiv:2302.11446v1 [math.AT] 22 Feb 2023

摘要

本文旨在提出一种简单的方法，以降低并控制随机矩阵的条件数，并研究对良态和病态矩阵所形成点云的持久同调（PH）的影响。对于使用高斯/均匀分布随机生成的方阵，SVD-Surgery 程序的步骤为：（1）计算其奇异值分解（SVD）；（2）通过将一组较小奇异值替换为该列表中各元素的凸线性组合，来替换对角因子；（3）通过逆向执行奇异值分解计算新矩阵。对矩阵应用 SVD 手术，通常会使其对角因子与输入矩阵的不同。已知随机方阵的空间分布与其条件数的分布相关。因此，持久同调（PH）研究着重比较 SVD 手术对由大规模随机生成的良态和病态矩阵构成的点云，以及由其逆矩阵构成的点云的影响。这项工作旨在通过控制卷积滤波器组的条件数，稳定深度学习（DL）训练对医学图像的影响，从而减少过拟合，并增强模型对可容忍图像噪声的鲁棒性。在训练期间将 SVD 手术应用于卷积滤波器，相当于对 DL 模型进行谱正则化，无需学习额外参数。我们将证明，对于若干由足够大规模卷积滤波器构成的点云，我们的简单策略能够保持滤波器范数，并根据所选的线性组合参数降低其逆矩阵的范数。此外，我们的方法显著改善了矩阵的良态性，并使其拓扑行为更加稳定。

关键词 病态性 · 正则化 · 条件数 · 奇异值分解 · 拓扑数据分析 · 代数拓扑 · 持久同调 · 随机高斯分布

1 引言

尽管深度学习模型在计算机视觉任务中取得了显著成功并不断进步，但人工智能在不同领域的部署仍面临严峻障碍：如何开发既稳健、又能在训练数据之外良好泛化的深度神经网络 [1]。精确而稳定的数值算法对于构建稳健可靠的计算模型至关重要 [2]。深度学习模型中的数值不稳定性，部分源于大量参数/超参数的使用，以及存在浮点误差和结果不准确问题的数据。对于卷积神经网络而言，其庞大的权重中反复执行反向传播算法，是造成不稳定性的一个明显因素；该算法通过控制梯度下降的增长，使模型性能适应不同的训练样本区域。本文关注对卷积神经网络训练所导致的各种权重矩阵条件数波动进行经验估计，并将其视为卷积层及整体模型性能不稳定性的潜在来源。我们将提出一种基于谱的方法，以减少并控制这种不必要的波动。

方阵 $n \times n$ 矩阵 A 作为线性变换 $\mathbb{R}^{n\times n} \to \mathbb{R}$ 时的条件数 $\kappa(A)$，用于衡量计算其作用时对输入数据扰动和舍入误差的敏感性，定义为在所有非零x上，$sup\|Ax\|/\|x\|$。它取决于其逆矩阵的计算受到下溢影响的程度

（即$det(A)$ 与 0 的差异有多显著）。A 的稳定作用意味着，输入数据的微小变化通常会导致输出数据的微小变化，而这些变化受条件数倒数的约束。因此，A 的条件数越高，其作用对微小数据扰动就越不稳定，这类矩阵被称为病态矩阵。事实上，随机矩阵的条件数分布直接描述了精度损失（以位数计），以及由于病态性导致在迭代求解线性方程组时的收敛速度， [3]。最初，矩阵的条件数由 A. Turing 于 [4]首次提出。此后，矩阵和数值问题的条件数在 [5–7]中得到了全面研究。计算 $\kappa(A)$ 最常见且高效稳定的方法，是计算A 的 SVD，并计算A的最大奇异值与最小非零奇异值之比， [8]。

J. W. Demmel 在 [6], 中研究了随机矩阵条件数概率分布的上界和下界，并表明，包括矩阵求逆、特征值问题和多项式零点求解在内的病态问题集合都具有共同的代数和几何结构。特别是，Demmel 表明，在矩阵求逆的情况下，一个矩阵距离不可逆矩阵集合越远，其条件数就越小。因此，随机矩阵在其定义域中的空间分布，可以反映其条件数的分布。这些结果清楚地证明了我们利用拓扑数据分析（TDA）工具研究随机矩阵点云条件数稳定性的方法具有可行性。一般而言，TDA 可用于在不具备数据先验知识的情况下，捕捉度量空间中点云复杂拓扑结构和几何结构的信息（详见 [9] ）。自 21 世纪初以来，应用拓扑学借助持久同调（PH）工具进入了一个新时代，用于研究高维数据集的整体和局部形状。由 PH 工具生成的持久图（PD）的各种向量化表示，编码了 CNN 模型卷积滤波器点云的局部几何和全局拓扑信息， [10]。在此，我们将尝试确定 SVD-Surgery 程序对 CNN 良态和病态卷积滤波器点云的持久图（PD）的影响。

贡献： 我们提出一种基于奇异值分解的矩阵手术（SVD 手术）技术，用于调整矩阵的条件数，适合稳定病态卷积滤波器在图像数据集点云上的作用。该技术通过奇异值分解将方阵分解，替换较小的奇异值，再利用所得的奇异值对角矩阵重构原矩阵。SVD 手术在保持输入矩阵范数的同时，降低其逆矩阵的范数。这意味着 SVD 手术会改变逆矩阵点云的持久同调。我们预计，对矩阵点云（及其逆矩阵的点云）进行持久同调分析，能够深入揭示图像分析深度学习模型的稳定性表现。

2 动机性挑战的背景

本文的根本动机源于我们在一项具有挑战性的研究中遇到的具体需求：如何“训练高效、精简的卷积神经网络模型，使其能够学习超声（或任何放射影像）中具有判别力的特征，从而支持临床诊断决策”。具体而言，所开发模型的预测必须能够抵御可容忍的数据扰动，并在未见过的数据上测试时不易受到过拟合的影响。

在机器学习和深度学习中，梯度消失或梯度爆炸以及收敛性差通常是由病态性问题导致的。克服病态性最常见的方法包括正则化、数据归一化、重新参数化、标准化和随机丢弃。在使用极大规模的“自然”图像数据集训练深度卷积神经网络时，卷积滤波器权重/条目会随机初始化，并在大量训练过程中通过多个图像批次和若干训练轮次不断调整；每个训练轮次结束时，反向传播过程都会更新滤波器条目，以提升性能。滤波器条目的频繁更新会导致其条件数出现不可忽略、甚至显著的波动和不稳定性，从而使受训 CNN 模型产生敏感性， [11, 12]。CNN 模型的敏感性表现为过拟合、对噪声的鲁棒性降低，以及容易受到对抗性攻击， [13]。

迁移学习是开发用于分析超声（或其他放射学）图像数据集的 CNN 模型时常用的方法：将现有 CNN 模型（在自然图像上训练）的预训练滤波器及其他模型权重用作重新训练的初始化参数。不幸的是，在小型非自然图像数据集上采用迁移学习模式时，条件数不稳定性会加剧，导致模型性能不佳并出现过拟合。

3 相关工作

深度学习 CNN 模型涉及多种参数，其复杂性主要由各卷积层中卷积滤波器组的元素以及全连接神经网络层的元素决定。在设计初始化策略时，这些参数的范数和／或方差是主要考量因素，目的是加快训练优化并提升机器学习和深度学习任务中的模型性能。目前，大多数流行的卷积神经网络架构都使用均值为零、方差受控且依赖或不依赖层的高斯分布来初始化这些权重。Krizhevsky 等人 [14], 使用 0.01 的固定标准差来初始化每一层的权重。为避免梯度呈指数级消失或增长，并兼顾激活函数的特性，*Glorot* [15], 或 *He* [16], 权重采用各层方差可控的方式进行初始化。对于 Glorot 初始化，初始化方差取决于输入和输出神经元的数量；而 He 初始化的方差则与其提出的参数化整流激活单元（PReLU）密切相关，该单元旨在提升模型拟合能力，同时降低过拟合风险。在上述所有初始化策略中，都没有明确考虑滤波器的条件数或其在训练期间的稳定性。我们的研究发现，在这些情况下，训练结束后几乎所有卷积滤波器都存在严重病态，因此会对其在非自然图像迁移学习中的应用产生不利影响。近期还有一些方法被提出，用于控制网络层的范数，例如 GradInit [17] 和 MetaInit [18]。这些方法能够加快收敛，同时提升模型性能和稳定性。然而，这两种方法都需要额外的可训练参数，而且无法保证训练期间的条件数得到控制。

近期，许多研究通过对可训练深度学习模型权重施加正交条件，探讨了与我们的目标密切相关的问题。这些研究包括正交归一和正交权重初始化技术、 [19–21], 正交卷积 [22], 正交正则化器 [23], 正交深度神经网络 [24], 以及正交权重归一化 [25]。鉴于正交／正交归一矩阵具有最优的良态性，这些研究间接支持了我们的假设，即深度学习模型过拟合与学习到的卷积滤波器的条件数有关。尽管这些研究并未明确讨论权重矩阵条件数的不稳定性，但它们符合一种新兴范式，即对神经网络层的权重矩阵进行谱正则化。例如，J. Wang 等人 [22], 指出，对卷积滤波器施加正交约束，是克服深度卷积神经网络模型训练不稳定性并提升性能的理想方法。此外，A. Sinha [23], 指出，病态的学习所得权重矩阵会增加神经网络受到对抗性攻击的易感性。事实上，他们的正交正则化旨在将学习所得权重矩阵的条件数保持在足够低的水平，并且在 MNIST 和 F-MNIST 自然图像数据集上测试时，展示了对抗准确率的提升。S. Li 等人在 [24],中指出，现有的谱正则化方案大多旨在提升实证应用中的训练效果；他们使用基于泛化误差（GE）概念的界限，对此类方法进行了理论分析。泛化误差以训练算法和应用特征空间的等距性为基础进行定义。他们得出结论：当深度神经网络的每个权重矩阵都具有相等的奇异值谱时，泛化误差的界限最优，并将此类模型称为正交深度神经网络。为克服严格正交深度神经网络的高计算开销，他们通过定期应用硬正则化方案——奇异值界定（SVB），定义了近似正交深度神经网络。总的来说，事实证明，在训练过程中控制权重的行为能够加快训练过程，并降低模型在训练集上过拟合的可能性，例如， [26], 权重标准化／重参数化 [27],中心化权重归一化 [28], 以及使用牛顿迭代实现可控正交化 [29]。上述大多数方法专为处理用于自然图像分析的可训练深度学习模型而开发，人们可能会认为，这些方法在训练过程中会频繁使用，例如在每个训练轮次／批次之后使用。然而，已知最先进的深度学习模型似乎都没有隐式地采用这些方法。事实上，我们对这些常用深度学习模型的调查发现，其最终卷积滤波器的条件数都很高，呈病态 [11]。

我们的文献综述发现，再调节和正则化长期以来一直用于分析应用，以减少/控制所述计算中的病态性。20 世纪 80 年代末，E. Rothwell 和 B. Drachman [30], 提出了一种迭代方法，通过对矩阵的非零奇异值进行正则化，来降低病态矩阵问题的条件数。每次迭代时，都会将正则化参数与奇异值的比值加到矩阵奇异值分解中的每个对角元素上。该算法不适用于我们的动机性挑战。此外，范数的变化取决于正则化参数。

近年来，人们对使用拓扑数据分析来分析各种类型和复杂度的数据集中的点云越来越感兴趣。例如，借助 PHtools，人们在捕捉高维数据集中的局部和全局拓扑特征与几何特征方面取得了显著进展并获得了深入认识，其中包括

传统方法， [31]。拓扑数据分析也已被用于解读深度学习和卷积神经网络在各层学习到的参数 [11,32,33],，以及将基于拓扑的方法融入深度学习 [34–39]。我们将使用拓扑数据分析评估良态和病态随机矩阵在奇异值分解手术前后，其矩阵／滤波器（及其逆矩阵）点云的空间分布。

4 拓扑数据分析

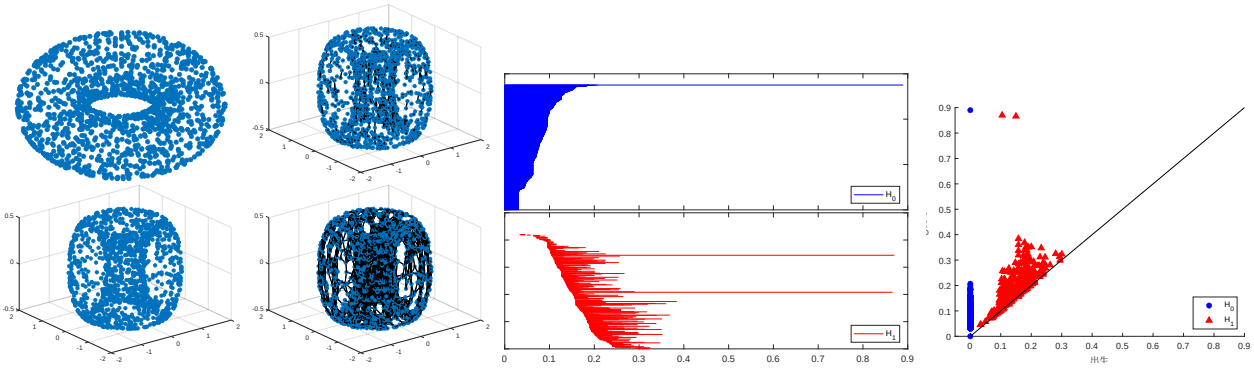
本节简要介绍持久同调的预备知识，并说明如何设置随机生成的矩阵点云，以研究其拓扑行为。

点云的持久同调：持久同调是拓扑数据分析（TDA）的一种计算工具，它通过记录逐步三角剖分形状的拓扑特征，来概括从度量空间采样得到的数据记录点云的空间分布。具体而言，按照一系列不断增大的距离/相似度阈值连接数据点对。对于点云 X 和一组递增阈值 $\{\alpha_i\}_0^m$，该 TDA 过程生成的形状 S(X) 是由单纯复形构成的序列 $\{\mathrm{S(X)}_i\}_0^m$，并按包含关系排序。Vietoris–Rips 单纯复形（VR）是构造 S(X) 最常用的方法，因为它简单易用，而 Ripser 是用于构造 VR 的工具， [40]。距离阈值序列称为 S(X) 的过滤。S(X) 的拓扑特征包括不同维度的孔洞或空腔数量，这些数量称为贝蒂数，分别对应于 $\{\mathrm{S(X)}_i\}_0^m$ 的各个组成部分。对于 $\mathrm{j} \leq 0$, 第 j 个贝蒂数 $\mathrm{B}j\mathrm{(S(X)}_i\mathrm{)}$ 分别通过计数以下对象获得：$\mathrm{B}_0 = \#$（连通分支）、$\mathrm{B}_1 = \#$（由多于 3 条边构成的空环）、$\mathrm{B}_2 = \#$（由多于 4 个面围成的三维空腔）等。注意，$\mathrm{B}j\mathrm{(S}_i\mathrm{(X))}$ 是单纯复形 $\mathrm{S}_i\mathrm{(X)}$ 的第 j 阶奇异同调的生成元集合。对于过滤 $\{\alpha_i\}_0^m$，对 X 进行拓扑数据分析时，依据的是 $\mathrm{B}j\mathrm{(S(X)}_i\mathrm{)}$ 中每个元素在 $\mathrm{i} \to \mathrm{m}$ 时的持久性。此处，每个元素的持久性定义为其出生（首次出现）与死亡（消失）之间的差值。通常用垂直堆叠的一组条形码直观表示 $\mathrm{B}j\mathrm{(S(X)}_i\mathrm{)}$，每个元素对应一条水平直线，连接其出生时刻与死亡时刻。更严格的详细描述见 [41–43]。为简便起见，$\mathrm{B}j\mathrm{(S}_i\mathrm{(X))}$ 的条形码集合和持久图（PD）统称为 $\mathrm{H}j$。

任意维度下，点云所产生的持久同调条形码的分析由
持久图(PD)由平面第一象限中点的多重集构成$(x = \text{出生}, y = \text{死亡})$，这些点位于直线$y = x$上方或线上。PD中的每个标记点都对应于给定维度的持久同调群的一个生成元，并由一对坐标(出生,死亡)表示。为了展示这些持久同调信息的可视化表示，我们在环面表面上随机采样了1500个点，构成一个点云：

$$T = \{(x,y,z) \in \mathbb{R}^3 : (\sqrt{x^2+y^2} - a)^2 + z^2 = b^2\}$$

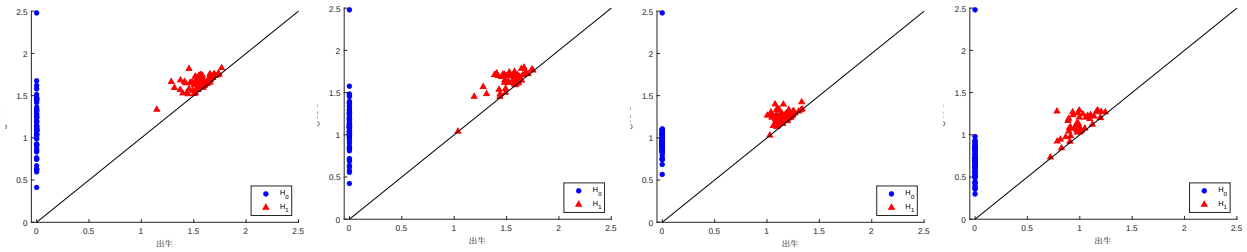
图1如下所示，展示了该点云，以及其在两个维度上的PH对应的条形码和PD表示。两个较长的 $1-$ 维度 持久条形码代表两个空圆盘，它们的笛卡尔积生成环面。这两个孔的持久长度取决于生成圆的半径$(a - b,b)$。在本例中，$a = 2b$。较短条形码集合的持久长度与点云大小成反比。噪声采样只会影响较短的条形码。



(a) 点云　(b) $d \leq 0.1$ & 0.2　(c) 持久条形码　(d) 持久图

图 1：点云示意图：(a) 环面上的点；(b) 连接距离不超过$d = 0.1$ 和 0.2 的邻近点，以及它们以持久条形码和持久图表示的拓扑结构。

我们之所以关注将PH研究与我们提出的矩阵手术联系起来，是因为滤波器的良态和病态点云之间存在显著的拓扑行为差异（通过PDs进行可视化）。图2展示了这一点：比较了两个随机生成的、由64个3x3矩阵组成的点云及其逆矩阵，并依据它们最低和最高的条件数进行了对比。



(a) 良态 $\mathcal{X}$ (b) 良态 $\mathcal{X}^{-1}$ (c) 病态 $\mathcal{X}$ (d) 病态 $\mathcal{X}^{-1}$

图 2：表示不同点云的点云的持久图，这些点云分别代表：(a) 良态矩阵，(b) (a) 中矩阵的逆矩阵，(c) 病态矩阵，(d) (c) 中矩阵的逆矩阵。

我们注意到，在两个维度上，良态矩阵与病态矩阵的持久性图之间存在差异，其各自逆矩阵的概率密度之间也存在差异。然而，当我们考察原始矩阵及其逆点云的持久性图时发现，在第 0 维中，与病态矩阵相比，良态矩阵的空间分布变化很小。我们提出的矩阵手术旨在控制输出矩阵的概率密度与其逆点云的概率密度之间的差异。

5 矩阵手术

本节将介绍执行矩阵手术的研究框架，其目标是降低并控制矩阵的条件数。假设矩阵 $A \in \mathbb{R}^{m\times n}$ 为非奇异矩阵，并基于随机高斯分布或均匀分布生成。A 的条件数定义如下：

$$\kappa(A) = \|A\|\|A^{-1}\| = \frac{\sigma_1}{\sigma_n}$$

其中 $\|.\|$ 是矩阵的范数，我们关注欧几里得范数（L_2-范数）； σ_1 和 σ_n 分别是 A 的最大和最小奇异值。如果输入的微小变化会导致输出发生巨大变化，则称矩阵是病态的；如果输入的微小变化只会导致输出发生相对较小的变化，则称矩阵是良态的。或者说，条件数较低（接近 1）的矩阵称为良态矩阵，而条件数较高的矩阵称为病态矩阵；正交矩阵的理想条件数为 1。接下来，由于条件数由最大和最小奇异值决定，我们将介绍一种基于奇异值矩阵修改奇异值分解的简单方法。回顾一下，方阵$A \in \mathbb{R}^{n\times n}$ 的奇异值分解定义如下：

$$A = U\Sigma V^T$$

其中$U \in \mathbb{R}^{m\times m}$ 和 $V \in \mathbb{R}^{n\times n}$ 是左右正交奇异向量（酉矩阵）；对角矩阵 $\Sigma = diag(\sigma_1, ..., \sigma_n) \in \mathbb{R}^{n\times n}$ 中的元素是奇异值，其中 $\Sigma = \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$。下文介绍的 SVD 手术同样适用于矩形矩阵。

5.1 基于奇异值分解的手术

广义而言，SVD 手术是指通过变换矩阵来改善其条件数的过程。具体来说，它针对正交性／标准正交性较差的矩阵，利用其左右正交奇异向量以及新的奇异值对角矩阵，将其替换为条件更良好的矩阵。根据输出矩阵需要满足的特性以适应具体用途，SVD 手术可以通过多种方式实现。对于任意矩阵 A，对 A 进行 SVD 手术后，会得到一个大小相同的新矩阵，如下所示：

1. 计算其奇异值分解，

2. 从对角矩阵因子 Σ 构造另一个对角矩阵 $\tilde{\Sigma}$，替换较小的奇异值，同时保持其降序排列

$$\tilde{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & ... & 0 \\ 0 & \tilde{\sigma}_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ... & \tilde{\sigma}_n \end{bmatrix}$$

其中，更新后的奇异值 $\tilde{\sigma}_i$'经过选择，以保持较低的条件数，同时确保新的对角元素仍按单调递减排列，并且

3. 按如下方式重构输出矩阵 $\tilde{A}$：

$$\tilde{A} = U\tilde{\Sigma}V^T$$

改变奇异值相当于重新缩放矩阵作用沿U 和 V的左、右正交向量方向产生的效应，而单调性要求则确保对各种缩放进行合理控制。尽管 [22]的正交正则化方案和 [24] 的 SVB 方案，在用于改善对自然图像训练的深度学习模型的过拟合控制时，确实能够降低条件数，但这些方案并不满足单调性。此外，将它们应用于超声图像数据集的深度学习训练时，其成效无法得到保证。SVB 在控制条件数方面比 SVD 手术严格得多，但并未分析这些矩阵及其逆矩阵的范数。我们提出的 SVD 手术专为所述应用场景而设计，旨在降低极高的条件数，同时保留原始矩阵的范数。将所有奇异值都替换为最大奇异值，会得到条件数为 1 的正交矩阵，但这相当于忽略或削弱训练数据沿某些奇异向量方向的显著变化所产生的影响（即降低学习效果）。在此，我们采用一种较为温和的策略来改变奇异值:

选择一个对角位置 $j, 1 < j < n$，令 $\tilde{\sigma}$ 为凸线性组合:

$$\tilde{\sigma}_k = \sum_{k=j-1}^{n-1} \alpha_k \sigma_k, \quad where \quad \alpha_i \geq 0, \quad \sum \alpha_i = 1$$

并为每个 $j \leq k \leq n$ 设置 $\tilde{\sigma}_k$。

j 的选择以及线性组合参数可以根据具体应用确定，也可以通过经验确定。

极端情况下，该策略包括将 $\tilde{\sigma} = \sigma_j$ 设为某值的可能性。与正交正则化策略相比，这种策略相当保守，因为它保持了奇异值的单调性。结合我们的应用背景，参数选择应根据层而定，但线性组合参数不应导致沿奇异向量方向对训练数据集特征进行显著重缩放。奇异值分解手术可应用于逆矩阵；然而，相同的替换策略和重构可能无法显著降低条件数。

示例 假设 B 是一个方阵，其中 $n = 3$ 从均值为 $\mu = 0$、标准差为 $\sigma = 0.01$ 的正态分布中抽取，如下所示:

$$B = \begin{bmatrix} -0.0196 & 0.0291 & -0.0106 \\ -0.0020 & 0.0083 & -0.0047 \\ -0.0121 & 0.0138 & -0.0027 \end{bmatrix}$$

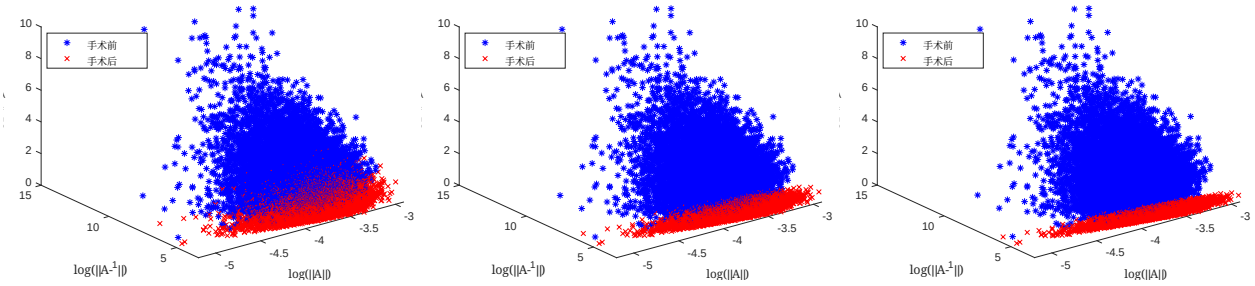
B 可用奇异值分解（SVD）表示并分解为:

$$U = \begin{bmatrix} -0.8745 & -0.1286 & -0.4676 \\ -0.2124 & -0.7652 & 0.6077 \\ -0.4359 & 0.6308 & 0.6419 \end{bmatrix} \Sigma = \begin{bmatrix} 0.0419 & 0.0000 & 0.0000 \\ 0.0000 & 0.0050 & 0.0000 \\ 0.0000 & 0.0000 & 0.0005 \end{bmatrix} V = \begin{bmatrix} 0.5452 & -0.7156 & 0.4367 \\ -0.7926 & -0.2703 & 0.5466 \\ 0.2731 & 0.6441 & 0.7145 \end{bmatrix}$$

U 和 V 是右、左正交奇异向量。B 的奇异值为 $\Sigma = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$，通过替换一个和/或两个奇异值，可以修改并重构 $\tilde{}$B1、 $\tilde{}$B$_2$, 和 $\tilde{}$B$_3$ ，分别使其成为 $\tilde{\Sigma}_1 = \text{diag}(\sigma_1, \sigma_2, \sigma_2)$、 $\tilde{\Sigma}_2 = \text{diag}(\sigma_1, \tilde{\sigma}_2, \tilde{\sigma}_3)$ 和 $\tilde{\Sigma}_3 = \text{diag}(\sigma_1, \sigma_1, \sigma_1)$。 $\tilde{\Sigma}_2$中的新奇异值是凸线性组合，满足 $\tilde{\sigma}_2 = 2\sigma_1/3 + \sigma_2/3$ 和 $\tilde{\sigma}_3 = \tilde{\sigma}_2$。

$$\tilde{B}_1 = \begin{bmatrix} -0 & & &

图3显示，经过手术处理后， 3×3 矩阵的逆矩阵范数分布发生了显著变化，条件数的分布也因此改变。线性组合能够根据奇异值的分布，将条件数范围控制在某个阈值以下。例如，图4中的三维示意图显示，通过将范围分别控制在 (b) 中的 3 以下和 (c) 中的 2 以下，条件数显著降低；其中 σ_2 和 σ_3 分别替换为 $\sigma_1/3 + 2\sigma_2/3$ 和 $(\sigma_1 + \sigma_2)/2$。两组矩阵经过矩阵手术后的新最小和最大条件数分别为 $[1.004, 2.687]$ 和 $[1.003, 1.88]$。



(a) $\tilde{\sigma}_3 = \sigma_2$ (b) $\tilde{\sigma}_2 = \tilde{\sigma}_3 = \sigma_1/3 + 2\sigma_2/3$ (c) $\tilde{\sigma}_2 = \tilde{\sigma}_3 = (\sigma_1 + \sigma_2)/2$

图 4：一组随机高斯矩阵在矩阵手术前后的示意图。图中显示了范数、逆矩阵的范数以及条件数（对数）。其中，在（a）中，最小奇异值 σ_3 被替换为 σ_2；在（b）和（b）中， σ_2 和 σ_3 被替换为由 σ_1 和 σ_2构成的新线性组合。

5.3 SVD 手术对卷积滤波器点云持久图的影响

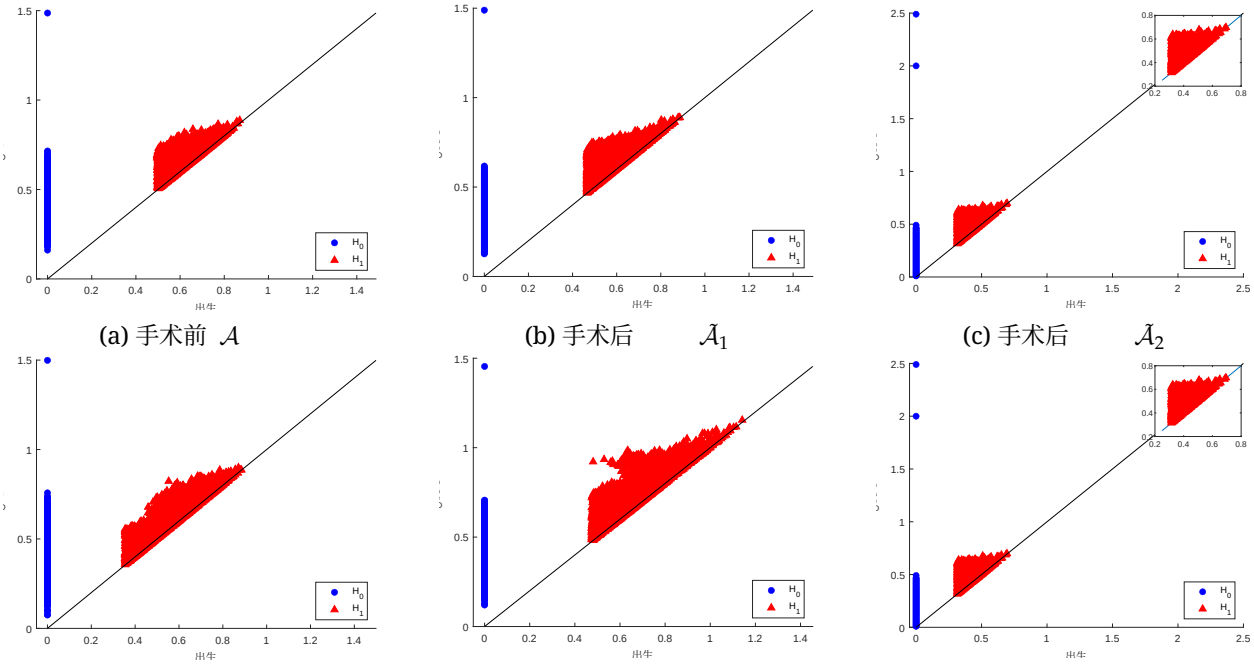
在训练期间控制 CNN 滤波器条件数的一种方法是控制其奇异值。SVD 手术的实现可以集成到定制的 CNN 模型中，用于分析自然图像数据集以及超声图像数据集，作为滤波器正则化项。它可以应用于从头开始训练时的滤波器初始化、迁移学习训练中的预训练滤波器，以及训练期间通过反向传播在每个批次/训练轮次后修改的滤波器。

本节将使用第4节中讨论的持久同调工具，研究一组矩阵作为点云时的拓扑行为。对于任意大小为$n \times n$ 的滤波器，我们首先生成一组随机高斯$n \times n$ 矩阵，然后将矩阵元素归一化并展平，从而得到位于 $\mathbb{S}^{n\times n-1}$上的一个点云，该点云位于$(n \times n - 1)$ − 球面上。我们通过计算逆矩阵、将其元素归一化并展平，在 $\mathbb{S}^{n\times n-1}$ 中构造第二个点云。这里，我们仅展示这一过程在特定的 3×3 矩阵点云上的应用，所用矩阵对应于较小的两个奇异值的两种不同线性组合。更大尺寸滤波器的一般情况将在第一作者正在撰写的博士论文中讨论。

下方的图5展示了点云（原始点云及其逆矩阵点云）的H_0 和H_1, 持久图，以及经过矩阵手术后的持久图，对应于以下线性组合：（1）将 σ_2 和 σ_3 都替换为 σ_1（即 $\kappa(A) = 1$），以及（2）将 σ_3 替换为 σ_1。第一行对应于奇异值分解对原始点云持久图的影响，第二行则对应于逆矩阵点云。

显然， $\tilde{\mathcal{A}}_2$ 和 $\tilde{\mathcal{A}}_2^{-1}$的持久性图是等价的，这反映出该变换生成了正交的、条件数达到最优的矩阵。请注意，在这种情况下，逆矩阵恰好是原矩阵的转置。点云 $\tilde{\mathcal{A}}_1$ 和 $\tilde{\mathcal{A}}_1^{-1}$由一组矩阵 $\tilde{A}_1$ 和 $\tilde{A}_1^{-1}$组成，其中最小奇异值 σ_3 被替换为 σ_2，条件数范围从 $[1.2, 10256]$ 到 $[1.006, 17.14]$。通常使用瓶颈距离等距离度量来确定持久性图之间的差异和相似性。

我们持续进行的研究表明，在以下情况下使用 SVD 手术处理卷积层滤波器，可显著提升卷积神经网络模型对可容忍数据扰动的鲁棒性，以及对未见数据的泛化能力：(1) 从头开始初始化时；(2) 使用预训练滤波器时；(3) 在训练批次和/或训练周期中；尤其是 (4) 将上述方法结合使用时。此外，在训练期间控制条件数可以稳定每个卷积层滤波器的拓扑行为。后续论文将探讨矩阵手术和拓扑数据分析在卷积神经网络中的这些应用。



(a) 手术前 $\mathcal{A}$　(b) 手术后 $\tilde{\mathcal{A}}_1$　(c) 手术后 $\tilde{\mathcal{A}}_2$

(d) 手术前 $\mathcal{A}^{-1}$　(e) 手术后 $\tilde{\mathcal{A}}_1^{-1}$　(f) 手术后 $\tilde{\mathcal{A}}_2^{-1}$

图5　：基于奇异值分解的手术前后点云 $\mathcal{A}$ 和 $\mathcal{A}^{-1}$ 的持久图。

6 结论

我们介绍了一种简单的矩阵手术策略，用于降低并控制$n \times n$矩阵的条件数：从输入矩阵中选定的对角元素开始，将其及其后所有奇异值替换为该列表的凸线性组合，再重构矩阵，从而大幅降低其条件数。我们证明，将这一策略应用于若干由足够大的卷积滤波器构成的点云时，可以保留滤波器的范数，并根据所选的线性组合参数降低其逆矩阵的范数。我们的方法显著改善了条件数和持久同调。针对我们的动机性挑战，SVD 手术方法能够在训练期间（从头开始训练或采用迁移学习模式）高效地控制卷积层滤波器的条件数，同时不会增加模型复杂度，也无需学习额外的超参数。该方法能够保持奇异值的单调性，并且可以根据层设置线性组合参数，以避免沿奇异向量对训练数据集特征进行大幅重新缩放。

致谢

本研究由 Ten-D AI Medical Technologies Ltd. 资助。

参考文献

[1] M. J. Colbrook, V. Antun, and A. C. Hansen，“稳定且准确地计算神经网络的困难：论深度学习的障碍与 Smale第18问题”，《美国国家科学院院刊》，第119卷，第12期，页码 e2107151119，2022年。[2] N. J. Higham，数值算法的准确性与稳定性。SIAM，2002年。[3] A. Edelman，特征值与条件数。博士论文，麻省理工学院，1989年。[4] A. M. Turing，“矩阵运算中的舍入误差”，《力学与应用数学季刊》，第1卷，第1期，页码287–308，1948年。[5] J. R. Rice，“条件数理论”，《*SIAM*数值分析杂志》，第3卷，第2期，页码287–310，1966年。[6] J. W. Demmel，“病态条件的几何学”，《复杂性杂志》，第3卷，第2期，页码201–229，1987年。

[7] D. J. Higham，“条件数及其条件数”，《线性代数及其应用》，第214卷，页码193–213，1995年。[8] V. Klema and A. Laub，“奇异值分解：计算方法及其若干应用”，IEEE《自动控制汇刊》，第25卷，第2期，页码164–176，1980年。[9] F. Chazal and B. Michel，“拓扑数据分析导论：面向数据科学家的基础与实践”，《人工智能前沿》，第4卷，2017年10月。[10] H. Adams and M. Moy，“应用于机器学习的拓扑：从全局到局部”，《人工智能前沿》，页码54，2021年。[11] J. Ghafuri, H. Du, and S. Jassim，“用于医学图像分析的CNN卷积层的拓扑特征”，载于《2020年移动多媒体/图像处理、安全与应用》，第11399卷，页码229–240，SPIE，2020年。[12] J. Ghafuri, H. Du, and S. Jassim，“预训练CNN滤波器的敏感性与稳定性”，载于《2021年多模态图像利用与学习》，第11734卷，页码79–89，SPIE，2021年。[13] I. J. Goodfellow, J. Shlens, and C. Szegedy，“解释并利用对抗样本”，arXiv预印本 arXiv:1412.6572，2014年。[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton，“使用深度卷积神经网络进行ImageNet分类”，《ACM通信》，第60卷，第6期，页码84–90，2017年。[15] X. Glorot and Y. Bengio，“理解深度前馈神经网络训练的困难”，载于《第十三届人工智能与统计学国际会议论文集》，第9卷，页码249–256，2010年。[16] K. He, X. Zhang, S. Ren, and J. Sun，“深入研究整流器：在ImageNet分类中超越人类水平性能”，载于《IEEE国际计算机视觉会议论文集》，第2015 Inter卷，页码1026–1034，2015年2月。[17] C. Zhu, R. Ni, Z. Xu, K. Kong, W. R. Huang, and T. Goldstein，“GradInit：学习如何初始化神经网络，以实现稳定高效的训练”，《神经信息处理系统进展》，第20卷，页码16410–16422，2021年2月。[18] Y. N. Dauphin and S. Schoenholz，“MetaInit：通过学习初始化来初始化学习”，《神经信息处理系统进展》，第32卷，2019年。[19] D. Xie, J. Xiong, and S. Pu，“你所需的不止是良好的初始化：探索结合标准正交性与调制、训练极深卷积神经网络的更优解”，载于《IEEE计算机视觉与模式识别会议论文集》，页码6176–6185，2017年3月。[20] D. Mishkin and J. Matas，“你所需的只是良好的初始化”，arXiv预印本 arXiv:1511.06422，2015年11月。[21] A. M. Saxe, J. L. McClelland, and S. Ganguli，“深度线性神经网络学习的非线性动力学的精确解”，arXiv预印本 arXiv:1312.6120，2013年12月。[22] J. Wang, Y. Chen, R. Chakraborty, and S. X. Yu，“正交卷积神经网络”，载于《IEEE/CVF计算机视觉与模式识别会议（CVPR）论文集》，2020年6月。[23] A. Sinha, M. Singh, B. Krishnamurthy, A. Sinha, B. Krishnamurthy, M. Singh, and B. Krishnamurthy，“对抗环境中的神经网络与病态权重空间”，载于《欧洲机器学习与数据库知识发现联合会议》，第11329卷 LNAI，页码177–190，Springer，Springer，2018年。[24] K. Jia, S. Li, Y. Wen, T. Liu, D. Tao, K. Jia, Y. Wen, T. Liu, and D. Tao，“正交深度神经网络”，IEEE《模式分析与机器智能汇刊》，第43卷，页码1352–1368，2021年4月。[25] L. Huang, X. Liu, B. Lang, A. Yu, Y. Wang, and B. Li，“正交权重归一化：深度神经网络中多个相互依赖的Stiefel流形上的优化问题的解”，载于《AAAI人工智能会议论文集》，第32卷，2018年。[26] S. Qiao, H. Wang, C. Liu, W. Shen, and A. Yuille，“结合批次-通道归一化和权重标准化的微批训练”，arXiv预印本 arXiv:1903.10520，2019年。[27] T. Salimans, D. P. Kingma, T. S. Openai, and D. P. K. Openai，“权重归一化：加速深度神经网络训练的一种简单重参数化方法”，《神经信息处理系统进展》，第29卷，页码901–909，2016年2月。

[28] L. Huang、X. Liu、Y. Liu、B. Lang和D. Tao，“用于加速深度神经网络训练的中心化权重归一化”，载于*IEEE*国际计算机视觉会议论文集，第2803–2811页，2017年。[29] L. Huang、L. Liu、F. Zhu、D. Wan、Z. Yuan、B. Li和L. Shao，“训练深度神经网络中的可控正交化”，*IEEE*计算机学会计算机视觉与模式识别会议论文集，第6428–6437页，2020年4月。[30] E. Rothwell和B. Drachman，“求解病态矩阵问题的统一方法”，工程数值方法国际期刊，第2卷，第609–620页，1989年3月。[31] R. Turkeš、G. Montúfar和N. Otter，“论持久同调的有效性”，*arXiv*预印本 *arXiv:2206.10551*，2022年。[32] R. B. Gabrielsson、G. Carlsson、R. Bruel Gabrielsson、G. Carlsson、R. B. Gabrielsson和G. Carlsson，“神经网络拓扑的阐释与解读”，载于*2019*年第*18*届*IEEE*国际机器学习与应用会议（*ICMLA*），第1069–1076页，IEEE，2019年。[33] G. Magai和A. Ayzenberg，“深度学习中数据流形的拓扑与几何”，*arXiv*预印本 *arXiv:2204.08624*，2022年。[34] C. Hofer、R. Kwitt、M. Niethammer和A. Uhl，“利用拓扑特征进行深度学习”，《神经信息处理系统进展》，第30卷，2017年。[35] B. Rieck、M. Togninalli、C. Bock、M. Moor、M. Horn、T. Gumbsch和K. Borgwardt，“神经持久性：利用代数拓扑度量深度神经网络的复杂度”，*arXiv*预印本 *arXiv:1812.09764*，2018年。[36] S. Ebli、M. Defferrard和G. Spreemann，“单纯形神经网络”，*arXiv*预印本 *arXiv:2010.03633*，2020年。[37] M. Hajij和K. Istvan，“深度学习的拓扑框架”，*arXiv*预印本 *arXiv:2008.13697*，2020年。[38] C. S. Hu、A. Lawson、J. S. Chen、Y. M. Chung、C. Smyth和S. M. Yang，“TopoResNet：一种混合深度学习架构及其在皮肤病变分类中的应用”，《数学》*2021*年第*9*卷第*2924*页，第9卷，第2924页，2021年11月。[39] R. Gonzalez-Diaz、M. A. Gutiérrez-Naranjo和E. Paluzo-Hidalgo，“基于拓扑的代表性数据集：减少神经网络训练资源”，《神经计算与应用》，第1–17页，2022年。[40] U. Bauer，“Ripser：Vietoris–Rips持久条形码的高效计算”，《应用与计算拓扑学杂志》，第5卷，第391–423页，2021年9月。[41] H. Edelsbrunner、D. Letscher和A. Zomorodian，“拓扑持久性与简化”，载于第*41*届计算机科学基础年度研讨会论文集，第454–463页，IEEE，2000年。[42] R. Ghrist，“条形码：数据的持久拓扑”，《美国数学学会通报》，第45卷，第1期，第61–75页，2008年。[43] N. Otter、M. A. Porter、U. Tillmann、P. Grindrod和H. A. Harrington，“持久同调计算路线图”，《*EPJ*数据科学》，第6卷，第1–38页，2017年。