线性回归在机器学习与深度学习中的地位

线性回归，尽管在概念上相对简单，却在机器学习乃至更广阔的深度学习领域中占据着核心地位。其持久的相关性源于其固有的可解释性以及作为更复杂统计和机器学习模型前身的关键作用。Christopher M. Bishop 和 Hugh Bishop 合著的《深度学习：基础与概念》[1] 在其教学框架中策略性地将线性模型引入早期章节。具体而言，“线性模型”在第一章作为“教程示例”被提出，并在第四章进一步阐述为“单层网络：回归”[5]。这种循序渐进的结构强调了线性回归的基础重要性，在深入探讨复杂的深度学习架构之前，它为核心概念的建立奠定了基础。

这种教学策略的深远意义在于，它提供了一个清晰的路径，使学习者能够从简单的、可解释的模型过渡到更复杂的系统。通过在直观的线性回归背景下掌握模型规范、参数估计和误差最小化等基本思想，学习者能够构建一个稳固的概念框架。这个框架随后被用于理解多层神经网络和其他深度学习架构的复杂性，其中这些基础原理被扩展、抽象和规模化。因此，这种早期引入充当了关键的概念支架，为学习更高级的主题做好了准备。这进一步表明，对高级人工智能和机器学习概念的深入且持久的理解，从根本上取决于对基础统计和数学模型的扎实掌握。它隐含地论证了深度学习并非一个孤立的领域，而是建立在数十年经典机器学习研究基础上的演进，线性回归正是这种连续性和基础依赖性的一个典型例证。

线性回归的这种基础作用，在 Christopher M. Bishop 的开创性著作《模式识别与机器学习》（PRML）中也得到了证实，该书专门用一个章节来阐述“线性回归模型”[7]，这进一步突显了这些原理在机器学习文献中一贯且持久的重要性。

学习背景与目的

本学习心得是“共读AI新圣经-深度学习”学习计划的重要组成部分，其重点是 Bishop 和 Bishop 所著《深度学习：基础与概念》中关于线性回归的内容。本次学习的主要目标是综合线性回归的理论基础、实践方法和评估技术。这项工作旨在展示对线性回归核心原理及其作为高级机器学习模型不可或缺的构建模块的全面理解。报告的呈现方式将严格遵循当代研究文献所体现的正式而精确的学术风格[9]。

2. 线性回归模型

模型定义与基本形式

线性回归是一种统计建模技术，旨在建立并量化一个因变量（通常表示为目标 t）与一个或多个自变量（称为输入 x）之间的关系。其主要目标是根据一个 D 维的输入变量向量

预测目标变量的连续值 [10]。

线性回归模型最基本的形式是将目标变量表示为输入变量的线性组合：$y(x,w)=w0+w1x1+\cdots+wDxD$。在此公式中，$w=(w1,\ldots,wD)T$ 代表回归系数或权重的向量，而 $w0$ 是截距项，通常被称为偏差参数 [10]。这种结构确保了模型的输出是其参数的线性函数。

线性模型与基函数

Bishop 强调的一个关键澄清是，“线性模型”的定义特征在于它们在参数 w 上是线性的，而非严格地在输入变量 x 上线性 [10]。这种区别对于扩展模型的表达能力至关重要。

为了使模型能够捕捉与输入之间更复杂的非线性关系，可以将线性组合应用于输入变量的固定非线性基函数 $\phi j(x)$。此时，模型形式变为：$y(x,w)=w0+\sum j=1M-1wj\phi j(x)$。通过约定定义一个额外的虚拟基函数 $\phi 0(x)=1$，这可以紧凑地写为 $y(x,w)=wT\phi(x)$，其中 $\phi(x)=(\phi 0(x),\phi 1(x),\ldots,\phi M-1(x))T$ [10]。

利用非线性基函数使得函数 $y(x,w)$ 能够成为输入向量 x 的非线性映射，同时关键地保持了参数 w 的线性。这种参数上的线性显著简化了模型的分析可处理性，尤其是在参数估计方面 [10]。

对“线性”模型中“线性”的理解，通常会让人联想到与原始输入特征的直线关系。然而，Bishop 明确强调其关键特性在于其在参数上的线性，这是一个至关重要的区别。这意味着即使模型的输出是原始输入的复杂非线性函数（通过基函数实现），由于对权重的线性依赖，底层的优化问题仍然是可处理的。这个概念是基础性的，因为它展示了模型如何在不牺牲线性代数分析优势的情况下捕捉非线性，为理解更复杂的模型如何隐式地学习非线性变换提供了桥梁。这种理解突出了模型的“线性”是指其参数形式，而不一定是其与原始输入的关系。这一原理在深度学习中得到了隐式扩展和规模化：虽然单个神经元执行线性变换（矩阵乘法），但层间非线性激活函数的应用使得整个深度网络能够学习从输入到输出的高度复杂、非线性映射。因此，线性回归中的基函数概念预示了深度神经网络中隐藏层作为复杂特征提取器的作用。

误差函数与目标

训练线性回归模型需要定义一个误差或损失函数。该函数量化了模型预测输出与训练数据集中实际观测目标值之间的差异。学习算法的基本目标是识别出在所有训练数据点上最小化此误差函数的最佳参数向量 w。

对于线性回归，最广泛采用的误差函数是平方误差和（SSE），它构成了普通最小二乘法

（OLS）准则的基础。该准则旨在最小化模型预测 y(xn,w) 与每个数据点 n 的真实目标值 tn 之间平方差的总和，并对所有 N 个训练观测值进行聚合：E(w)=21∑n=1N{y(xn,w)−tn}2。乘以 1/2 是为了简化导数计算 [11]。

3. 参数估计：最小二乘法

最小二乘法的推导

为了确定最小化平方误差和的最优参数向量 w，优化中的标准方法是计算误差函数 E(w) 对 w 的梯度并将其设为零。此过程为 w 提供了闭合形式的分析解。

对于表示为 y(x,w)=wTϕ(x) 的线性模型（其中 ϕ(x) 是基函数向量，为偏置项增加了 ϕ0(x)=1），OLS 解由以下公式给出：w^=(ΦTΦ)−1ΦTt。其中，Φ 是设计矩阵，其每一行对应于给定数据点 xn 的基函数评估 ϕ(xn)，而 t 是目标值向量 [8]。

几何解释

从几何角度来看，普通最小二乘法解可以直观地理解为观测目标向量 t 在设计矩阵 Φ 中由基函数向量（或最简单情况下的输入向量）张成的列空间上的正交投影。预测值向量 t^=Φw^ 代表该子空间内与实际目标向量 t 在欧几里得距离上最接近的点。最小化此欧几里得距离正是平方误差和所实现的目标 [8]。

最大似然估计的视角

OLS 的一个强大统计解释是，它对应于模型参数的最大似然估计（MLE），前提是目标变量 t 是由确定性函数 y(x,w) 生成并受到附加高斯噪声 ϵ 干扰的。具体而言，t=y(x,w)+ϵ，其中噪声 ϵ 被假定为独立同分布（i.i.d.）并服从高斯分布，即 ϵ ~ N(0,σ2) [11]。

在此概率假设下，最大化观测数据关于参数 w 的似然函数，在数学上等价于最小化平方误差和。这种最大似然估计的视角为 OLS 提供了坚实的统计基础，并作为通向更高级概率模型（包括贝叶斯线性回归）的自然桥梁 [11]。

普通最小二乘法（一个代数最小化问题）与最大似然估计（一种概率推断方法）在高斯噪声假设下的数学等价性，是一个深刻且基础性的理解。这种联系将 OLS 从单纯的曲线拟合技术提升为一种具有统计基础的方法。它意味着在最小二乘意义上的“最佳拟合”也是在给定常见噪声假设下最可能的拟合。这种概率框架至关重要，因为它构成了理解许多高级机器学习技术的基础，包括深度学习中各种损失函数的设计（例如，用于分类的交叉熵

，它也是在不同分布假设（如伯努利或多项式）下从 MLE 推导出来的）。这种理解在经验误差最小化和严格统计推断之间建立了直接而强大的联系。它展示了即使是像线性回归这样看似简单的模型，也可以从多个互补的角度（代数、几何和概率）进行理解，从而丰富了学习者的理论基础。这种多方面的理解对于设计、解释和调试更复杂的深度学习模型是不可或缺的，因为损失函数几乎总是植根于概率原理（例如，负对数似然）。

4. 线性回归的假设与局限

核心假设

为了使线性回归模型推导出的统计推断（如置信区间、p 值和假设检验）在统计上有效和可靠，必须满足关于误差项（残差）的几个关键假设。虽然 Bishop 和 Bishop 的《深度学习：基础与概念》中可能没有明确列举这些假设，但 Christopher M. Bishop 的《模式识别与机器学习》 [14] 广泛强调了概率论对于全面理解机器学习的必要性，这暗示了这些统计先决条件。权威的外部资源，如宾夕法尼亚州立大学的 STAT 500 [15] 和 Kaggle [16]，提供了这些关键假设的全面概述：

线性性 (Linearity)： 自变量与因变量之间的关系必须是线性的。这意味着所选择的模型正确地指定了关系的潜在函数形式 [15]。违反此假设会导致参数估计有偏差和模型拟合不良。

误差独立性 (Independence of Errors)： 与每个观测值相关的残差（误差）必须相互独立。此假设在时间序列数据或空间数据中尤为关键，因为自相关（误差项之间的相关性）可能频繁发生，导致标准误差被低估和统计显著性被夸大 [15]。

误差正态性 (Normality of Errors)： 残差必须近似服从正态分布。虽然对于足够大的样本量，中心极限定理可以减轻非正态性的影响，但此假设是统计检验和置信区间有效性的正式要求 [15]。

同方差性 (Homoscedasticity)： 残差的方差在自变量的所有水平上应保持不变。异方差性（误差方差不相等）会导致参数估计效率低下和标准误差有偏差，从而影响假设检验的可靠性 [15]。

无完美多重共线性 (No Perfect Multicollinearity)： 自变量之间不应存在完美的线性关系。完美的共线性会使设计矩阵 $\Phi^T\Phi$ 变为奇异矩阵，从而无法唯一确定回归系数 [16]。高但非完美的共线性会导致参数估计不稳定且高度敏感。

这些统计假设并非抽象的理论细节；它们是模型推断和预测有效性和可靠性的直接先决条件。例如，如果违反了同方差性假设，回归系数的标准误差将不正确，导致 p 值和置信区间出现错误。因此，理解这些假设对于有效的模型诊断、识别误差来源以及就模型适用性或替代方法的需要做出明智决策至关重要。

模型局限性分析

尽管线性回归具有分析可处理性和可解释性，但其仍存在固有的局限性。正如 Bishop 所指出的，最简单的线性回归形式（模型在输入变量上严格线性）对其捕捉复杂关系的能力施加了“显著限制” [10]。即使通过基函数进行增强，模型在参数上仍然是线性的，这限制了其在没有细致且通常依赖领域知识的特征工程的情况下，建模高度复杂、非线性交互的能力。

此外，违反上述核心假设会严重损害参数估计的可靠性以及从模型中得出的统计推断的有效性。例如，如果变量之间真实的潜在关系是高度非线性的，并且没有选择或学习到一套合适的基函数，那么线性模型将遭受高偏差，导致数据系统性地欠拟合。

在线性回归中使用非线性基函数 [10] 来捕捉复杂输入-输出关系，是深度神经网络架构的一个概念性前身。在线性回归中，这些变换是由模型设计者明确定义和选择的。相比之下，深度神经网络通过其多层结构和非线性激活函数，从数据中隐式地学习这些复杂、分层的特征变换。Bishop 对简单线性模型“显著限制”的观察 [10] 强调了这种变换的内在必要性，以有效建模真实世界数据中复杂的模式。这种从显式特征工程到自动化特征学习的演变，是经典机器学习向深度学习转变的一个决定性特征。这一观察建立了清晰的概念谱系，展示了如何将建模非线性关系这一根本挑战，最初通过经典线性回归中精心设计的基函数来解决，演变为定义深度学习的复杂、多层、非线性架构。这种演进说明了从手动、专家驱动的特征工程到数据驱动、自动化特征学习的范式转变，这是深度神经网络成功背后的核心优势和驱动力之一。

5. 模型评估与正则化

常用评估指标

对回归模型性能进行严格评估对于理解其有效性、诊断其弱点以及促进不同模型之间的比较至关重要。关键指标用于定量衡量模型预测值与实际观测值之间的差异 [17]。

均方误差 (Mean Squared Error, MSE)： 该指标计算预测值 (y^i) 与实际目标值 (yi) 之间平方差的平均值：MSE=N1∑i=1N(yi−y^i)2。由于平方运算，MSE 会严重惩罚较大的误差，使其对异常值敏感。其可微分性也使其成为优化中广泛适用的损失函数选择 [17]。

均方根误差 (Root Mean Squared Error, RMSE)： RMSE 定义为 MSE 的平方根：RMSE=MSE。这种转换使误差指标的单位与因变量相同，显著增强了其可解释性。与 MSE 类似，RMSE 也强调了较大误差的影响 [17]。

平均绝对误差 (Mean Absolute Error, MAE)： MAE 计算预测值与实际值之间绝对差

的平均值：MAE=N1∑i=1N|yi−y^i|。MAE 对异常值不如 MSE 或 RMSE 敏感，并且它保留了输出变量的原始单位，提供了直接且直观的可解释性 [17]。

R 平方 (R2, Coefficient of Determination)： 该指标量化了因变量中可由自变量解释的方差比例。其计算公式为 R2=1−SSTSSR，其中 SSR 是残差平方和，SST 是总平方和。R2 的值范围从 0 到 1，值越高表示模型对观测数据的拟合越好。它常被称为“拟合优度”的度量 [17]。

过拟合与正则化概念

过拟合 (Overfitting)： 机器学习模型训练中普遍存在的挑战是过拟合。当模型过于细致地学习训练数据时，就会发生这种现象，它不仅捕捉了潜在的模式，还捕捉了训练集中特有的随机噪声和特定特性。因此，过拟合模型在训练数据上表现出色，但在未见数据上泛化能力差。当模型相对于训练数据的大小或固有噪声水平过于复杂时，这种情况尤其常见。Bishop 的 PRML [19] 承认“过拟合现象”是机器学习中的一个基本概念。

正则化 (Regularization)： 为了对抗过拟合并增强模型的泛化能力，需要采用正则化技术。这些方法在模型的原始误差函数中引入一个惩罚项，通过限制模型参数（权重）的大小来有效阻止模型过度复杂。这鼓励了更简单的模型，从而减少了记忆训练数据的倾向。

过拟合和正则化的概念与机器学习中基本的偏差-方差权衡密切相关。过拟合是高方差（模型对特定训练数据过度敏感，导致对未见数据的预测波动较大）与低偏差（模型非常紧密地拟合训练数据，甚至是噪声）的体现。正则化 [11] 作为一种机制，向模型引入受控的偏差量。通过限制模型的复杂性（例如，通过收缩权重），它有效地降低了方差，从而提高了模型对新的、未见数据的泛化能力。这种微妙的平衡对于实现最佳预测性能至关重要，并且是所有机器学习范式（包括深度学习）中模型选择、超参数调整和算法设计中反复出现的主题。这一观察强调，追求最佳模型性能不仅仅是最小化训练数据上的误差。相反，它是在模型准确表示训练数据（最小化偏差）的能力和在新、未见数据上表现良好（最小化方差）的能力之间进行复杂的平衡。这一原则深刻地影响着深度学习中的架构决策、优化策略以及各种正则化技术（例如，Dropout、批量归一化）的应用，因为高度复杂的模型本质上容易受到高方差和过拟合的影响。

误差函数（例如，MSE、MAE）是学习算法在训练阶段旨在最小化的主要定量目标。正则化项 [11] 然后在数学上附加到这些核心目标上。这种增强将目标函数转换为一个复合实体，它同时寻求最小化预测与实际值之间的差异，并对模型的复杂性施加约束（例如，通过惩罚大权重）。优化过程随后最小化这个组合目标函数。这个概念框架可以直接转移，并构成了深度学习的基础，其中各种损失函数（例如，用于分类的交叉熵，各种回归损失）被优化，并经常通过 L1 或 L2 正则化（通常称为“权重衰减”）进行增强，以提高泛化能力并防止过拟合。这展示了理论机器学习原理（拟合数据和控制模型复杂性）到实

际优化问题的优雅数学转换。它强调了将这些双重目标公式化为一个可以算法最小化的单一统一函数的强大能力。这种设计原则在统计建模和机器学习中无处不在，为理解模型如何学习以及如何管理其复杂性提供了连贯的框架。

6. 总结与展望

学习心得总结

本次对线性回归的全面探索，在 Christopher M. Bishop 的基础著作（《深度学习：基础与概念》和《模式识别与机器学习》）的精心指导下，深刻地强化了其作为远不止一个基本统计工具的地位。它是一个理解预测建模核心机制的基本范式。我们系统地涵盖了其正式定义，阐明了通过普通最小二乘法进行参数估计的优雅过程及其通过最大似然估计的强大概率解释，批判性地审视了支撑其有效性的基本统计假设，并讨论了模型评估和正则化的关键策略。

Bishop 著作 [1] 在呈现这些复杂概念时的教学清晰度，经常通过直观的几何和严谨的概率解释 [8] 来丰富，为任何有抱负的机器学习从业者或研究人员提供了宝贵的概念工具包。

线性回归作为深度学习基础的意义

线性回归中固有的基本原理——即定义数学模型、制定可量化的误差函数、通过迭代方法（如梯度下降，在概念上与 [8] 中的顺序学习相关）或闭合形式解优化参数、理解潜在的模型假设以及通过正则化缓解过拟合——可以直接转移并扩展到深度神经网络领域。

此外，核心概念如损失函数的设计（例如，深度学习中回归任务的均方误差）、优化算法的迭代性质（例如，随机梯度下降，它是线性模型中基于梯度的优化的一种扩展），以及正则化技术（例如，L2 正则化/权重衰减，甚至像 Dropout 这样的概念，可以看作是一种模型平均形式，如 [5] 中所述）的应用，都是在线性回归背景下首次遇到并理解的原理的直接后代和高级阐述。

线性回归，尽管其表面上看似简单，却可以作为理解支撑更复杂机器学习模型（包括深度神经网络）核心原理的“微观世界”。它清晰地概括了基本阶段：定义模型结构、通过目标函数量化预测误差、优化模型参数以最小化此误差、理解验证模型的统计假设以及采用正则化来增强泛化能力。这些元素正是深度学习模型的基本构建块，尽管它们被放大和抽象化了。整个深度网络随后有效地从数据中隐式地学习了一组复杂的、多层的“基函数”。因此，对线性回归的扎实掌握为剖析和理解深度学习的机制提供了宝贵的概念和实践模板。对线性回归的强大基础理解，使从业者和研究人员具备了分析工具、问题解决框架和批判性思维能力，这些都是在日益复杂的深度学习领域中有效调试、解释和创新所必需的。

参考文献

1. Deep Learning: Foundations and Concepts (Hardcover) - Prairie Lights Books, 访问时间为 六月 21, 2025, https://www.prairielights.com/book/9783031454677
2. Deep Learning - Foundations and Concepts, 访问时间为 六月 21, 2025, https://www.bishopbook.com/
3. Deep Learning: Foundations and Concepts - Microsoft Research, 访问时间为 六月 21, 2025, https://www.microsoft.com/en-us/research/publication/deep-learning-foundations-and-concepts/
4. Deep Learning: Foundations and Concepts (2024) - Porchlight Book Company, 访问时间为 六月 21, 2025, https://www.porchlightbooks.com/products/deep-learning-christopher-m-bishop-9783031454677
5. Deep Learning: Foundations and Concepts 9783031454677, 9783031454684 - DOKUMEN.PUB, 访问时间为 六月 21, 2025, https://dokumen.pub/deep-learning-foundations-and-concepts-9783031454677-9783031454684.html
6. Deep Learning | springerprofessional.de, 访问时间为 六月 21, 2025, https://www.springerprofessional.de/en/deep-learning/26251802
7. Pattern Recognition and Machine Learning by Christopher M. Bishop, Hardcover, 访问时间为 六月 21, 2025, https://www.barnesandnoble.com/w/pattern-recognition-and-machine-learning-christopher-m-bishop/1100527382
8. Pattern Recognition and Machine Learning [Hardcover ed.] 0387310738, 9780387310732 - DOKUMEN.PUB, 访问时间为 六月 21, 2025, https://dokumen.pub/pattern-recognition-and-machine-learning-hardcover-nbsped-0387310738-9780387310732.html
9. 2502.18036v3.pdf
10. Bishop C. Deep Learning Foundations and Concepts. Sec. 4.1.1 A ..., 访问时间为 六月 21, 2025, https://www.scribd.com/document/770596753/Bishop-C-Deep-Learning-Foundations-and-Concepts-Sec-4-1-1-a-4-1-3-Pag-112-a-118
11. Solutions to “Pattern Recognition and Machine Learning” by Bishop - tommyodland.com, 访问时间为 六月 21, 2025, https://tommyodland.com/files/edu/bishop_solutions.pdf
12. Bishop's PRML, Chapter 4, 访问时间为 六月 21, 2025, https://www.di.fc.ul.pt/~jpn/r/PRML/chapter4.html
13. PRML Chapter 3 | PPT - SlideShare, 访问时间为 六月 21, 2025, https://www.slideshare.net/slideshow/prml-chapter-3-249686763/249686763
14. Pattern Recognition and Machine Learning - Microsoft, 访问时间为 六月 21,

2025, https://www.microsoft.com/en-us/research/wp-content/uploads/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf
15. 9.2.3 - Assumptions for the SLR Model | STAT 500, 访问时间为 六月 21, 2025, https://online.stat.psu.edu/stat500/lesson/9/9.2/9.2.3
16. Step by Step Assumptions - Linear Regression - Kaggle, 访问时间为 六月 21, 2025, https://www.kaggle.com/code/shrutimechlearn/step-by-step-assumptions-linear-regression
17. 10 Regression Metrics For Machine Learning & Practical How To Guide - Spot Intelligence, 访问时间为 六月 21, 2025, https://spotintelligence.com/2024/03/27/regression-metrics-for-machine-learning/
18. Know The Best Evaluation Metrics for Your Regression Model - Analytics Vidhya, 访问时间为 六月 21, 2025, https://www.analyticsvidhya.com/blog/2021/05/know-the-best-evaluation-metrics-for-your-regression-model/
19. Book Review: Pattern Recognition and Machine Learning - SPIE Digital Library, 访问时间为 六月 21, 2025, https://www.spiedigitallibrary.org/journals/journal-of-electronic-imaging/volume-16/issue-04/049901/Book-Review-Pattern-Recognition-and-Machine-Learning/10.1117/1.2819119.full