磐石计划 01 · B 卷

命题报告（学习路径设计）

版本 v3.0；每卷 21 题、150 分；对象为本次对话中已经接触基础数学语言与深度学习入门内容的学习者。现无实际答卷与追踪数据，本文不报告实测难度、区分度、平均分或“保证掌握率”。

一、任务从筛选改为学习

保留正式试卷外形和四类题型，但题目不是互不相关的检查点。A 卷侧重恢复定义与生成方法，B 卷将同一方法放到新的约束、对象或证明任务中。首次作答的错误用于选择支持，不用于给学习者贴能力标签。

版式参照用户提供的 2025 新课标 I 卷数学 PDF 的居中标题、连续题号、题型说明、普通选项和页脚；解析参照用户提供的 2024 模拟答案册的【答案】【分析】【详解】【点睛】。保留 21 题，不照抄参考卷 19 题结构、原数学题或机构署名。参考文件为版式与解析组织来源，不是本套深度学习理论的依据。

题型	题数	分值	学习功能
单项选择	8	40	对象、定义、短计算
多项选择	3	18	理由、反例与概念分离
填空	4	20	从条件重建表达式
解答	6	72	分步发现、证明与迁移

二、核心知识回收网络

知识链	卷间回收	完成后的可见证据
符号与函数族	A1→B1→A19	能区分集合、元素、函数、函数值及不同参数的同一函数。
矩阵、向量与展平	A3/A4→B3/B4→B14→B20	能给每个对象标形状，并区分计算结果与存储表示。
复合、激活与表达	A2/A8/A12→B2/B16→A20→B17	能从分段函数和几何结构说明非线性的作用。
监督学习与风险	A5/A13→A17→B13→B19	能从记录构造损失，再将经验与总体、固定与选后保证分开。
导数、反传与优化	B5→A18→B8→B18→B21	能写局部导数、完整链式因子与条件化下降证明。
范数、积分与误差	A7/A11/A14→B7/B11→B20	能用不同范数评价误差，证明不等式并构造尖峰反例。
可辨识性与压缩	B1→A19→B12/B14→B20	分别讨论函数唯一、参数唯一、存储数量、近似误差。
稳定性、归纳偏置与算力	A10/A15→B10→A21→B21	知道等变不等于不变、残差不等于收缩、MAC 不等于延迟。

箭头表示建议回访，不表示后题必须依赖前题的答案；正式题面提供必要定义，避免前题做错就无法进入后题。学习顺序可按 A 卷小段订正后进入 B 卷对应题，而非先硬做完两卷。

三、逐题命题说明

1. 不同参数可能是同一个函数

已学基础：讲义第 1、9 章。要生成的证据：能解释参数到函数的多对一映射，并检查定义域。

设计动作：模型族的元素是函数，不是参数标签；要比较函数在整个给定输入集合的值。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把参数组数当函数数；把笛卡尔积当并集；只在 x=0 比较导致所有函数看似一样。再使用位置：A1、A19。

2. ReLU 零区间与复合不等式

已学基础：讲义第 3 章。要生成的证据：能从分段定义推导集合，而不是只求一个根。

设计动作：先做复合，再把激活等于零转成输入不大于零的不等式。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把 ReLU=0 解成内部 =0；把闭端点漏掉。再使用位置：A2、B16、B17。

3. 长乘积的维度审计

已学基础：讲义第 2 章。要生成的证据：能在多步乘积中持续跟踪维度，并用转置恒等式重组计算。

设计动作：先在每项旁标行列，不需要知道矩阵元素，就能排除没有定义的式子。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：只凭字母熟悉程度判断；以为所有写在一起的矩阵都能相乘。再使用位置：A3、A18、B20。

4. 逐元素运算与范数的组合

已学基础：讲义第 2、7 章。要生成的证据：能跨内积、外积、逐元素乘积与范数保持对象类型清楚。

设计动作：先逐位置相乘得到完整向量，再对该向量取范数。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：先求内积 1 再取绝对值；认为乘 0 的坐标可以改变维数。再使用位置：A3、A7、A18。

5. 导数的每个因子来自哪里

已学基础：讲义第 4、5 章。要生成的证据：能解释导数数值和链式因子，并回到一次真实参数更新。

设计动作：固定样本后只剩 w 可变。先把模型输出代入损失，再求复合导数。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：将 y 或 x 也当作正在求导的变量；漏链式因子；导数与损失本身混淆。再使用位置：A13、A18、B18。

6. 偏置决定边界，约定决定端点

已学基础：讲义第 3 章。要生成的证据：能把阈值、偏置、不等式和几何边界互相转换。

设计动作：先算分数，再按照本题零点约定分类；边界只由分数等于零确定。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：沿用其他数学语境的 sign(0)=0；把边界本身判为负类；忽略偏置。再使用位置：A16、B16。

7. 尖峰误差让不同范数分离

已学基础：讲义第 7 章。要生成的证据：能构造而非只口述平均误差小、最坏误差大的反例。

设计动作：每个 n 下仍是一个矩形脉冲：高度 n，宽度 n 的负二次方；用这一几何结构计算，不逐点积分。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：认为一种范数趋零就意味着所有范数趋零；只看高度不看宽度。再使用位置：A11、A14、

B17。

8. 负梯度方向与有限步长

已学基础：讲义第 6 章。要生成的证据：能用精确余项或精确递推判断步长，不依赖口号。

设计动作：一次更新分三步：旧点梯度、参数改变量、新点的真实损失；不能拿一阶近似替代最后一步。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：看到负梯度就断言必下降；只算参数不验证真实损失；混用近似号与等号。再使用位置：A13、B18。

9. 监督信号、反传与优化器

已学基础：讲义第 4、5、12 章。要生成的证据：能按训练信号来源和计算步骤分类，而不靠名词记忆。

设计动作：检查“训练目标从哪里来”和“这一步计算什么”，避免以有无人工类别划分所有学习。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把所有标签都当人工类别；把反向传播当成学习率选择算法。再使用位置：A9、A18。

10. 用具体运算检验等变与不变

已学基础：讲义第 12 章。要生成的证据：能亲算等变等式、反例和聚合不变，并认识其代价。

设计动作：直接计算移位前后结果。把响应的坐标重排与响应的数值保持不变分开。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：仅因名字是卷积就写不变；因果掩码被解读为因果发现。再使用位置：A10、B21。

11. 用反例检查四种看似合理的结论

已学基础：讲义第 6、7、11 章。要生成的证据：能区分全称证明和反例，并说清范数使用条件。

设计动作：正确的全称不等式需要证明，错误的全称命题只需一个符合条件的反例。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：检验几个数都成立就说全称成立；用不符合假设的反例；误称 $q<1$ 也满足所有范数公理。再使用位置：A6、A7、B18、B21。

12. 二次交叉系数与对称约束

已学基础：讲义第 9 章。要生成的证据：能辨明数据条件和额外结构约束的不同作用。

设计动作：将矩阵模型展开成单项式，再比较同类项系数。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：直接把 a 和 b 都设为 6；认为 x1x2 与 x2x1 独立；把对称约束当数据已经给出。再使用位置：A19、B20。

13. 分布不同，也可能对某函数有同一期望

已学基础：讲义第 8 章。要生成的证据：能避免“分布不同所以期望必不同”的逻辑跳跃。

设计动作：先按重复次数构造经验概率，再算目标函数在每种取值上的值。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：拿真实概率 1/2 代替经验频率；认为只要分布不同，任何期望都不可能相同。再使用位置：A5、A17、B19。

14. 参数矩阵展平不是更换模型

已学基础：讲义第 2、9 章。要生成的证据：能从网络结构计算参数量，并区分形状、数量与有效自由度。

设计动作：每一层权重的行数等于输出维数，列数等于输入维数；偏置跟输出同维。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把行列反过来；漏偏置；把参数向量当输出向量；把展平误认为线性模型。再使用位置：A4、A12、B20。

15. 约束最优不一定是驻点

已学基础：讲义第 1、6 章。要生成的证据：能在有约束和最小值不取到的情形中使用正确优化记号。

设计动作：先检查允许的参数范围，再看目标函数在这个区间往哪个方向减小。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：直接解 $L'=0$ 得到不合法点；把最小值与最小化点集合颠倒。再使用位置：A6、B18。

16. 非线性表征使 XOR 可分

已学基础：讲义第 3 章。要生成的证据：能从不可能性到构造再到充要条件，理解特征表示的作用。

设计动作：先证明原空间中不可能，再构造新特征。这样知道非线性到底改变了什么，而不是只见到一次成功拟合。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：尝试很多参数当成不可能性证明；负类边界写成非严格；构造参数只验证必要性不验证充分性。再使用位置：A12、A16、A20。

17. 从折点结构构造平方逼近

已学基础：讲义第 10 章。要生成的证据：能用插值与状态复用重建显式逼近，而不是只背 O(log) 结论。

设计动作：A20 制造折点，这题给它们一个任务：修正平方函数插值的误差。每轮比较旧端点、新中点，再用分段仿射的唯一性补齐整个区间。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把粗网格宽度与细网格宽度混淆；漏误差中的 1/4；重算所有三角波导致 $O(m^2)$；把平方特例说成完整通用逼近。再使用位置：A20、A14、B7。

18. 从局部下降到全程函数值收敛

已学基础：讲义第 6 章。要生成的证据：能独立给出收敛证明并明确它证明的是函数值、不是所有参数性质。

设计动作：两条条件各有分工：光滑上界控制走一步的代价；PL 条件将梯度大小转为距最优值的差。要把它们连起来才能得到递推。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：把充分步长范围写成所有函数必要范围；把线性收敛理解成误差一次函数；从函数值收敛直接推出参数唯一。再使用位置：A6、A13、B8。

19. 从固定模型界到训练后选择的保证

已学基础：讲义第 8 章。要生成的证据：能说明统一事件为何覆盖数据依赖选择，并把优化误差加入界。

设计动作：训练会挑模型，所以不能事后假装模型预先固定。先保证所有候选模型同时受控，再允许在这个事件里任意挑选。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：将置信度当单条预测正确率；忘记 J 代价；把随机选择的模型直接代入固定模型界；风险差中漏一个 ε。再使用位置：A9、A17、B13。

20. 低秩近似要同时有下界与可达构造

已学基础：讲义第 9 章。要生成的证据：能证明最佳性，并把压缩误差转为带输入条件的输出误差。

设计动作：不在四个冗余因子参数里蛮算。低秩矩阵会丢失某个方向；沿这个方向看原矩阵仍保留多少，再构造最好的保留方案。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：找到一个误差为 1 的矩阵就宣布最优；把最优矩阵与唯一因子混淆；把矩阵误差不加条件等同于任务误差。再使用位置：A19、A21、B14。

21. 残差、梯度与数值误差的三重辨析

已学基础：讲义第 5、11 章。要生成的证据：能用同一结构比较前向误差与反向导数，给出条件化而非名称式解释。

设计动作：同一条链，既可以逐层相乘导数，也可先把状态相邻差求和再求导；比较两种写法后，检查“+1” 有没有被其他项抵消。支持先后：H1 指出观察对象；H2 给中间关系；解析给完整证明；最后撤去提示完成同题改动。

防止的假掌握：遗漏 dxk/dx0；见到 +1 就断言梯度 ≥ 1；将数值误差有界当梯度不消失。再使用位置：A18、A21、B11。

四、经典论文不只作背景装饰

经典证明均以新定义题的方式出现：题面补足超出已学内容的前提，学生完成可由现有工具推出的关键部分。不能将原定理答案直接当作可引用结论；若题目要求证明次数界、收敛率或最佳性，须写出中间不等式及依据。

论文来源与改编边界

以下只列本册使用的来源；原文编号与教学题编号不同。正文重新给出全部考查假设。基础恒等式不冒充论文专属定理，教学特例不替代原论文一般结论。

【P2】 Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. (1986). Learning representations by back-propagating errors. Nature 323, 533–536.

原文定位 / 教学范围：公式 (1)–(6)：前向计算、平方损失及链式求导。

DOI / 标识：10.1038/323533a0

【P5】 Yarotsky, D. (2017). Error bounds for approximations with deep ReLU networks. Neural Networks 94, 103–114.

原文定位 / 教学范围：第 3.1 节、Proposition 2：平方函数的深 ReLU 逼近；采用 arXiv v3 的命题编号。

DOI / 标识：10.1016/j.neunet.2017.07.002

【P6】 Karimi, H.; Nutini, J.; Schmidt, M. (2016). Linear Convergence of Gradient and Proximal-Gradient Methods Under the Polyak–Łojasiewicz Condition. ECML PKDD, 795–811.

原文定位 / 教学范围：第 2 节、Theorem 1：光滑性与 PL 条件下的梯度下降收敛。

DOI / 标识：10.1007/978-3-319-46128-1_50

【P7】 Hoeffding, W. (1963). Probability Inequalities for Sums of Bounded Random Variables. Journal of the American Statistical Association 58(301), 13–30.

原文定位 / 教学范围：有界独立变量的集中界；试卷将双侧均值界作为已知工具。

DOI / 标识：10.1080/01621459.1963.10500830

【P8】 Vapnik, V. N.; Chervonenkis, A. Ya. (1971). On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities. Theory of Probability and Its Applications 16(2), 264–280.

原文定位 / 教学范围：一致收敛背景；试卷只证有限模型类版本，不涉及完整 VC 维理论。

DOI / 标识：10.1137/1116025

【P9】 Eckart, C.; Young, G. (1936). The Approximation of One Matrix by Another of Lower Rank. Psychometrika 1, 211–218.

原文定位 / 教学范围：最佳低秩近似经典结果；试卷只证二维对角矩阵的秩一特例。

DOI / 标识：10.1007/BF02288367

【P10】 He, K.; Zhang, X.; Ren, S.; Sun, J. (2016). Identity Mappings in Deep Residual Networks. ECCV, 630–645.

原文定位 / 教学范围：第 2 节公式 (3)–(5)：恒等捷径的前向展开与反向传播；试卷采用标量可微特例。

DOI / 标识：10.1007/978-3-319-46493-0_38

【P11】 Vaswani, A., et al. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems 30.

原文定位 / 教学范围：编码器与解码器注意力、位置与因果掩码背景，不在本卷考完整注意力推导。

DOI / 标识：arXiv:1706.03762

【P12】 Goodfellow, I. J., et al. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems 27.

原文定位 / 教学范围：判别器学习“数据 / 生成”来源，不要求人工任务类别。

DOI / 标识：arXiv:1406.2661

五、学习设计依据与使用边界

闭卷重建与延迟再写的设计参考 Karpicke 与 Roediger（2008）关于提取练习的实验；同题更换条件后的迁移参考 Butler（2010）的实验问题。前者使用外语词汇，后者研究事实与概念保持、迁移，均不直接证明本套数学卷对某个人的学习效果。本卷的具体题目安排和下述 1 日 / 3 日 / 7 日回访属于教学设计，不是论文验证的最优时间表。

Karpicke, J. D.; Roediger, H. L. (2008). The Critical Importance of Retrieval for Learning. Science 319, 966–968. DOI: 10.1126/science.1152408。

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. Journal of Experimental Psychology: Learning, Memory, and Cognition 36(5), 1118–1133. DOI: 10.1037/a0019902。