

人工智能与线性代数课程深度融合的教学改革研究

——基于实证数据的成效验证

吴嘉润, 赵娜娜*, 苑孟理想, 唐子航, 刘新红

北京石油化工学院, 北京 102617

DOI:10.61369/ASDS.2026060006

摘 要 : 为探索人工智能技术与线性代数课程深度融合的有效路径, 本研究以某高校2024级511名学生为研究对象, 其中实验班101人实施AI融合教学干预(包含可视化工具辅助、AI智能答疑与自适应练习), 其余学生构成全级对照基准。通过期末统一考试成绩, 运用独立样本t检验、Mann-Whitney U检验、Kolmogorov-Smirnov双样本检验及卡方检验等多元统计方法验证教学成效。结果表明: 实验班平均分(78.71 ± 15.18)显著高于全级对照组(70.94 ± 18.62), 差异具有极显著统计学意义($t=3.939, p<0.001$, Cohen's $d=0.458$); 实验班优秀率(40.6%)较对照组(16.4%)提升24.2个百分点, 不及格率由27.2%降至16.8%。各题型成绩整体优于对照组, 提升最明显的是综合应用类题目。研究表明, AI辅助教学能够在一定程度上改善线性代数课程的学习成效, 也为后续教学改革提供了实证依据。

关 键 词 : 人工智能; 线性代数; 教学改革; 实证研究; 效应量分析

Research on Teaching Reform of Deep Integration of Artificial Intelligence and Linear Algebra Curriculum—Effectiveness Verification Based on Empirical Data

Wu Jiarun, Zhao Nana*, Yuan Menglixiang, Tang Zihang, Liu Xinhong

Beijing Institute of Petrochemical Technology, Beijing 102617

Abstract : To explore effective pathways for the deep integration of artificial intelligence (AI) technology with the linear algebra curriculum, this study enrolled 511 students from the 2024 cohort of a Chinese university. Among them, 101 students in the experimental class received AI-enhanced instruction—including visualization tools, intelligent Q&A systems, and adaptive practice platforms—while the remaining students served as the control group. Final examination scores were analyzed using independent-samples t-test, Mann-Whitney U test, two-sample Kolmogorov-Smirnov test, and chi-square test. Results show that the experimental class achieved a significantly higher mean score (78.71 ± 15.18) compared with the control group (70.94 ± 18.62) ($t=3.939, p<0.001$, Cohen's $d=0.458$). The excellent rate (≥ 90 points) of the experimental class (40.6%) exceeded that of the control group (16.4%) by 24.2 percentage points, while the failure rate dropped from 27.2% to 16.8%. Performance was better across all question types, especially in comprehensive application problems. These findings suggest that AI-assisted teaching can improve learning outcomes in linear algebra and offer a practical reference for teaching reform.

Keywords : artificial intelligence; linear algebra; teaching reform; empirical research; effect size analysis

引言

线性代数作为理工科专业的核心公共基础课程, 以其高度抽象性与强逻辑依赖性著称, 历来是高等数学教学改革的难点与焦点。传统教学模式以板书演算为主, 难以有效帮助学生建立矩阵变换、特征值等抽象概念的几何直觉, 学生普遍反映知识碎片化严重、“知其然而不知其所以然”。

随着人工智能技术进入教育场景, 课程教学的组织方式也在发生变化。教育部相关政策多次强调, 要推进AI与教学深度融合, 探

基金项目: 国家级大学生创新创业训练计划(2026J00115); 北京市高等教育学会课题(MS2025130); 北京市大学生创新创业训练计划(2026J00091); 北京市教委科研计划一般项目(KM202410017004)。

作者简介:

吴嘉润, 本科生, 研究方向: 数据挖掘与个性化推荐算法;

苑孟理想, 本科生, 研究方向: 人工智能与智能数据分析;

唐子航, 本科生, 研究方向: 数理统计与数据建模;

刘新红, 副教授, 博士, 研究方向: 应用数学。

通讯作者: 赵娜娜, 讲师, 博士, 研究方向: 概率论与数理统计。

素更加注重学生体验和学习效果的智慧教育模式^[1]。从已有研究看，知识图谱、自适应平台等工具已经在高等数学教学中出现了不少应用^[2-4]，但落到线性代数课程上，系统性的实证研究仍然不多，尤其是同时使用对照组设计、多种统计检验和效应量分析的研究更少。

基于这一背景，本研究以某高校2024级511名学生为样本，构建准实验研究框架，在统一考试评价体系下比较实验班（n=101）与全级对照基准的成绩表现，并结合参数与非参数方法检验差异是否稳定。我们希望通过真实教学数据，尽量回答一个更具体的问题：AI融合教学到底能否在这门课程中带来可观察、可验证的改进。

一、AI 融合教学模式设计

（一）设计理念

本研究把 AI 定位为教学辅助工具，而不是替代教师讲授的主体。它的作用不是直接给出结论，而是通过结构化呈现、即时反馈和个性化练习，帮助学生降低理解门槛、形成直观认识，并逐步提高知识迁移能力。这个思路主要借鉴了建构主义学习理论和认知负荷理论：前者强调学生要在有意义的情境中主动建构知识，后者提醒教学设计要尽量减少无效信息对工作记忆的占用^[5、6]。

据此，本研究将教学目标概括为三点：第一，帮助学生把矩阵运算、行列式、向量空间、线性变换、特征值与特征向量等核心概念学得更扎实；第二，提高综合应用题的解题质量，尤其是跨知识点的迁移能力；第三，根据学习过程中的反馈及时调整教学安排，让课堂更有针对性。

（二）AI 融入教学方式

本研究将 AI 嵌入教学过程，主要分为三个层次，彼此之间是递进关系。

第一层是可视化辅助理解。借助可视化工具，把矩阵变换、特征向量方向等抽象内容转化成动态几何演示。比如在讲矩阵乘法时，直接展示它对二维向量的作用效果，让学生把代数运算和几何变化联系起来。这样的呈现方式通常更容易帮助学生建立直观印象，也能缓解抽象概念带来的理解压力^[7]。

第二层是 AI 答疑和个性化练习。课后如果学生在某一步推导上卡住，可以借助 AI 工具查看分步解释，而不是只拿到一个最终答案。与此同时，系统会根据学生过往答题情况推送不同难度和不同侧重点的练习题，重点补足薄弱环节。

第三层是基于数据的教学调整。教师定期查看平台汇总的作业和测验结果，了解哪些知识点是全班普遍薄弱的，再据此调整课堂讲解顺序和训练重点。

（三）教学实施步骤

教学实施分为课前、课中、课后三个环节，具体如下：

课前，教师在每章开始前推送少量基础测试题，用来了解学生对核心知识点的掌握情况。AI 自动生成班级学情报告，教师据此判断本节课应该重点讲哪些内容；学生则可以结合自己的报告提前预习，把时间更多放在薄弱点上。

课中，教师结合动态可视化界面讲解知识之间的联系，并根据 AI 统计出来的高频错误点进行重点突破。课堂练习后，系统会即时反馈学生对各知识点的掌握情况，教师可以据此继续调整节奏。

课后，AI 根据课堂测验数据为每位学生生成差异化补救方案，推送相关微课和同类习题。与此同时，教学中适当加入数学史和跨专业应用案例，如计算机图形学、数据降维等，有助于提高学生兴趣，也能让他们看到课程内容的实际用途。

二、研究设计与方法

（一）研究对象与数据

本研究采用准实验设计，以某高校一个年级的理工科学生为对象。实验班共 101 人，接受为期一学期的 AI 融合教学干预；全级 511 人（包含实验班）作为参照基准，其中非实验班学生采用传统教学方式。两组学生使用相同教材和教学大纲，参加同一期末统一考试，评分标准一致，因此数据具有较好的可比性。

期末考试满分 100 分，题型结构如下：选择题 30 分（10 题 × 3 分），填空题 20 分（10 题 × 2 分），大题 50 分（第 3 至第 8 大题，前 4 题满分 10 分，后 2 题满分 5 分）。

（二）研究假设

本研究提出以下三项假设：

1. 接受 AI 融合教学的实验班期末总分显著高于全级对照基准。
2. 实验班在选择、填空和大题三个题型上的得分率均优于对照组，且综合应用题的提升最明显。
3. 教学干预的综合效应量达到中等及以上水平（Cohen's $d \geq 0.3$ ）。

（三）统计分析方法

为尽量提高结论的稳健性，本文从参数检验和非参数检验两个角度同时展开分析。

- ① 独立样本 t 检验——检验两组总分均值差异，适用于样本量较大、近似正态的情形；
- ② Mann-Whitney U 检验——非参数稳健性验证，不依赖正态分布假设，应对数据偏态情况；
- ③ Kolmogorov-Smirnov 双样本检验——检验两组成绩累积分布函数的整体形态差异；
- ④ χ^2 检验——比较优秀率、不及格率等离散指标的组间差异；
- ⑤ Cohen's d 效应量——量化干预的实质意义（0.2= 小效应，0.5= 中效应，0.8= 大效应）^[8]；
- ⑥ 各题型分组 t 检验 + 森林图——细粒度分析各题型干预效果及置信区间。

三、实证分析结果

（一）总体成绩比较

实验班平均分为78.71分（SD=15.18），全级对照组平均分为70.94分（SD=18.62），组间差值为7.77分。独立样本t检验结果显示，差异具有极显著统计学意义（ $t=3.939$ ， $df=610$ ， $p<0.001$ ）。由于Shapiro-Wilk检验显示两组数据均存在一定程度的非正态性，同时采用Mann-WhitneyU检验进行稳健性验证，结论一致（ $U=31815$ ， $p<0.001$ ）。效应量Cohen's $d=0.458$ ，达到中等效应水平，表明教学干预具有实质性的成效，而非仅具统计显著性（见表1）。

表1：描述统计指标对比

统计指标	全级对照组 (n=511)	实验班 (n=101)	差值
均值 ± 标准差	70.94 ± 18.62	78.71 ± 15.18	+7.77
中位数 (IQR)	74 (59 - 86)	80 (67 - 91)	+6
优秀率 (≥90分)	16.4%	40.6%	+24.2%
不及格率 (<60分)	27.2%	16.8%	-10.4%
最高分 / 最低分	100/15	100/19	—

注：优秀率与不及格率组间差异均经 χ^2 检验验证，均具有统计显著性（ $p<0.05$ ）。

成绩分布的频率直方图与核密度估计（图1）直观展示了两组的分布形态差异：实验班分布整体右移，峰值区域向高分段偏移，分布更为集中（标准差较对照组小3.44分）；对照组分布则呈现明显的左偏特征，低分段占比更高。

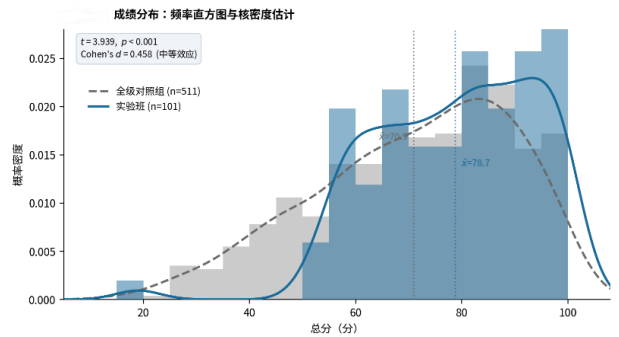


图1：成绩分布对比（频率直方图与核密度估计）

Figure 1 Score Distribution Comparison (Histogram and Kernel Density Estimation)

（二）成绩分布与等级结构

从等级结构分析（图2），实验班在优秀（≥90分）档次的占比为40.6%，较全级对照组（16.4%）提升24.2个百分点，差异经 χ^2 检验高度显著（ $\chi^2=26.74$ ， $p<0.001$ ）；“不及格”（<60分）档次由对照组的27.2%降至实验班的16.8%（ $\chi^2=5.21$ ， $p=0.022$ ）。值得注意的是，实验班的“良好”档次（75-89分）占比（30.7%）略低于对照组（31.9%），与之相对的是优秀率的大幅提升，提示AI融合教学对高分段学生的拉升效果尤为突出。

累积分布函数（CDF）分析进一步说明了两组之间的差别。Kolmogorov-Smirnov 双样本检验显示，两组CDF存在显著整

体偏移（ $D=0.208$ ， $p<0.001$ ）。从实际图形上看，实验班在各个分数节点上的累积百分比都更低，说明其整体成绩表现优于对照组，而不是只在某一小段分数上占优。

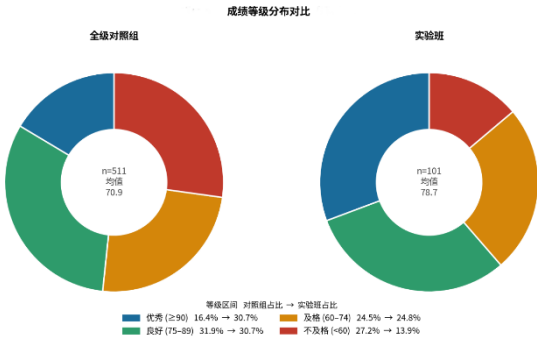


图2：成绩等级分布对比

Figure 2: Grade Distribution Comparison

（三）各题型得分率分析

按题型细看，实验班在8个题型上的得分率均高于对照组，说明AI融合教学对不同知识点都有一定帮助。

选择题和填空题方面，实验班得分率分别为84.1%和86.6%，高于对照组的79.1%和82.4%，两项差异均达到统计显著（ $p<0.05$ ）。相比之下，这类基础题的提升幅度不算特别大（分别提升5.0和4.2个百分点），可能是因为这部分内容更依赖日常练习和熟练度积累。

大题，也就是综合应用题，提升最明显。其中第7题提升23.5个百分点（ $p<0.001$ ），第8题提升13.3个百分点（ $p<0.001$ ），第4题提升12.1个百分点（ $p<0.001$ ）。这几道题都需要把多个知识模块连起来使用，例如线性变换、特征值分解和方程组求解等，说明AI融合教学对知识整合和迁移的帮助更突出。

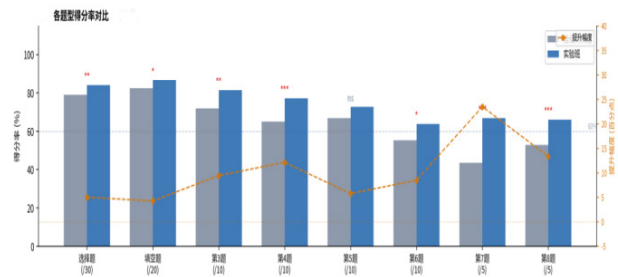


图3：各题型得分率对比与提升幅度

Figure 3: Section Score Rate Comparison and Improvement Margin

（四）效应量异质性分析

为进一步量化各题型的干预效果，计算各题型的Cohen's d 值并绘制森林图（图4）。结果显示，效应量存在显著的异质性：第7题效应量最大（ $d \approx 0.95$ ），远超其他题型；第4、8题次之（ d 分别约为0.60和0.55）；选择题和第5题效应量相对较小（ $d<0.3$ ），且第5题95%置信区间包含0值，提示该题型的组间差异在统计上尚不稳健。综合效应量 $d=0.458$ （95%CI 不含0），对应中等效应量水平，证实研究假设H3成立。

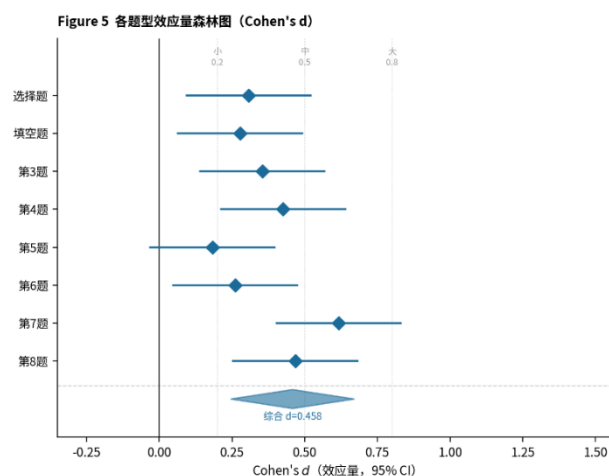


图4：各题型效应量森林图（Cohen's d, 95% 置信区间）

Figure4: ForestPlotofEffectSizesbyQuestionType(Cohen's d,95%CI)

（五）综合统计检验汇总

表2汇总了本研究所采用的全部统计检验结果，可见各方法均指向一致结论：AI 融合教学干预对线性代数成绩的提升效果具有极显著统计学意义，且在参数与非参数两类检验框架下均稳健成立，有效排除了结论依赖于分布假设的风险。

表2：假设检验综合结果汇总

检验方法	统计量	自由度 / 样本量	p 值	结论
独立样本 t 检验	$t=3.939$	$df=610$	<0.001	***
Mann-WhitneyU	$U=31815$	$n1=101,n2=511$	<0.001	***
K-S 双样本检验	$D=0.208$	—	<0.001	***
优秀率 χ^2 检验	$\chi^2=26.74$	$df=1$	<0.001	***
不及格率 χ^2 检验	$\chi^2=5.21$	$df=1$	0.022	*
综合效应量	$Cohen'sd=0.458$	中等效应	—	—

注：*** 表示 $p<0.001$ ，* 表示 $p<0.05$ ；Cohen's d 效应量解释标准：0.2= 小效应，0.5= 中效应，0.8= 大效应。

四、讨论

（一）结果解释

总体上看，本文结果与已有 AI 辅助教学研究的结论大体一致，即 AI 干预通常能够带来中等水平的正向效果^[9, 10]，但本文进一步看到，这种效果并不是平均铺开在所有题型上的，而是集中体现在综合应用题上。换句话说，AI 融合教学的优势更多在于帮助学生把零散知识串起来、用起来，而不仅仅是把记忆题做得更快。这一点也与 Mayer 关于多媒体学习的观点相吻合：恰当的视

觉呈现和符号呈现，有助于学生进入更深层的认知加工^[7]。

从成绩结构上看，实验班不及格率从 27.2% 降到 16.8%，优秀率从 16.4% 升到 40.6%，这两项变化是同时发生的。也就是说，AI 融合教学一方面帮助了原本基础薄弱的学生，另一方面也让成绩较好的学生有了进一步提升的空间，这和差异化教学中“分层支持”的思路比较一致。

（二）研究局限性

本研究也有以下几点局限，结论在解读时需要保持谨慎：

① 本研究是准实验设计，不是随机分组，因此实验班和对照组之间可能本来就存在学习动机等方面的差异，这一点很难完全排除。后续研究可以考虑随机分班，或者借助倾向得分匹配（PSM）尽量平衡组间差异。

② 本文主要使用期末考试成绩作为结果变量，没有充分覆盖学习过程中的变化，比如元认知策略、课堂投入度等。后续如果能加入过程性评价，结论会更完整。

③ 样本来源只有一所学校，外部推广时还需要更多学校、更多专业类型的验证，才能判断结论是否具有更强的普适性。

④ 教师因素没有完全控制。实验班和对照组如果由不同教师授课，教师风格本身也可能影响结果，这一部分在当前框架下不容易拆分出来。

五、结束语

本研究基于 511 名学生的真实考试数据，借助多种统计方法验证了人工智能融合教学在线性代数课程中的实际效果。研究结果支持三项核心假设：实验班总分显著优于全级对照基准（H1 成立）；各题型得分率整体提升，综合应用题最为突出（H2 成立）；综合效应量达到中等水平（H3 成立）。

总体来看，以 AI 辅助教学为主线的融合教学模式，确实能够帮助学生更好地理解线性代数中的核心概念，也能在一定程度上提升他们解决综合题的能力。它对不同层次学生都有积极作用，说明这种模式不只适合“学得快”的学生，也能对基础较弱的学生起到补充和支持作用。本文所采用的准实验设计和统计分析路径，也可以为同类课程改革研究提供一个可借鉴的样本。后续研究可继续从随机化设计、长期追踪和多校验证三个方向推进，以进一步增强因果推断的说服力。

参考文献

- [1] 教育部等九部门关于加快推进教育数字化的意见 [J]. 中华人民共和国教育部公报, 2025, (04): 16-20.
- [2] 杨文霞, 王卫华, 何朗, 等. 知识图谱赋能智慧教育的研究与实践——以武汉理工大学“线性代数”课程为例 [J]. 高等教育教育研究, 2023 (6): 111-117.
- [3] 刘媛媛, 高小伟. 基于 AI 的混合式教学模式优化——以“线性代数”课程为例 [J]. 大学教育, 2025 (32): 72-75.
- [4] 涂建华, 肖碧怡, 姜广峰. 构建微积分知识图谱助推一流课程建设 [J]. 中国大学教学, 2020 (11): 33-37.
- [5] 龙珠, 陈祥伟, 董波. 工科数学实验教学体系融合思政的研究与探索 [J]. 实验室研究与探索, 2023, 42 (11): 216-219.
- [6] Sweller, J. Cognitive load theory, learning difficulty, and instructional design [J]. Learning and Instruction, 1994, 4(4): 295-312.
- [7] Mayer, R.E. Multimedia Learning (3rd ed.) [M]. Cambridge University Press, 2020.
- [8] Cohen, J. Statistical Power Analysis for the Behavioral Sciences (2nd ed.) [M]. Lawrence Erlbaum Associates, 1988.
- [9] Kulik, J.A., & Fletcher, J.D. Effectiveness of intelligent tutoring systems: A meta-analytic review [J]. Review of Educational Research, 2016, 86(1): 42-78.
- [10] Hattie, J. Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement [M]. Routledge, 2009.