矩阵初等变换方法的计算稳定性与误差分析研究

黄佐赫
杭州师范大学

摘　要：矩阵初等变换法是数值线性代数中解线性方程组、求逆矩阵、算特征值的重要工具，它的计算稳定性和误差影响直接关系到算法的精度和可靠性。本文从理论和实验两个方面系统地研究矩阵初等变换法的计算稳定性以及误差传播规律。通过数值实验来验证理论分析的正确性，并给出在保证精度的前提下减小误差放大的办法。

关键词：矩阵初等变换；计算稳定性；误差分析；高斯消元；数值算法

随着科学计算和工程应用对数值精度要求的不断提高，矩阵运算方法的稳定性和可靠性成了研究热点。矩阵初等变换是线性代数课程中矩阵的一种重要且基础的运算法则，它是研究矩阵、向量组线性相关性、线性方程组的解、二次型以及线性空间等内容不可替代的工具。等 [1]。在计算机浮点环境中，舍入误差和病态问题常常导致计算结果偏离理论值，影响算法的稳定性和精度。传统上多把研究重点放在算法效率上，对于误差传播机制、稳定性控制的系统分析比较欠缺。通过对比分析不同的变换策略和主元的选择对误差放大有影响的因素，发现误差放大的主要原因，给出相应的改进方法，为高精度数值计算提供参考。

1 矩阵初等变换方法基础

1.1 矩阵的基本概念与运算性质

矩阵是数或者符号按行和列形式组成的矩形数组， 记作 $A = 【a_{ij}】_{m\times n}$， 其中 a_{ij} 表示第 i 行第 j 列的元素。矩阵是线性代数的一个基本工具，矩阵的运算有加法、数乘和矩阵乘法，当行数与列数相等时称为方阵。单位矩阵 I_n 的主对角线上的元素是 1，其余元素是 0，满足 $AI_n = I_nA = A$。矩阵乘法具有结合性、分配性，但不具有交换性。若存在矩阵 B 使得 $AB = BA = I$，则称 A 可逆，且 $B=A^{-1}$。矩阵的秩 rank(A) 描述了行向量或列向量的线性独立性，是研究初等变换及其稳定性的基本量 [2]。

1.2 初等行变换与列变换的定义

矩阵的初等行变换的三类是：(1) 将矩阵两行进行交 换，如将第 m 行与第 n 行交换；(2) 用一个不为零的数去乘矩阵中某一行的所有元素；(3) 用矩阵中某一行的所有元素。分别乘一个不为零的数，然后加到另外一行对应的元素中。

列变换形式与行变换完全对应，执行行变换相当于左乘初等矩阵 E，即 $A' = EA$；列变换为右乘初等矩阵 F，即 $A'' = AF$。若将第二行加上第一行的两倍，可令 E = [[1,0,0],[2,1,0],[0,0,1]]，则 $A` = EA$ 表示新的矩阵形式。

2 矩阵初等变换在计算中的应用

2.1 高斯消元法中的初等变换

高斯消元法是初等行变换在数值计算中的典型体现。对于线性方程组 $AX = b$，一系列行变换可将矩阵（可逆、顺序主子式大于 0） A 化为上三角矩阵 U，同时对向量 b 施加相应变换得到 b'。运算可表示为 $E_kE_{k-1}\cdots E_1A = U$，$E_kE_{k-1}\cdots E_1b = b'$。将各步操作综合后，可得 $A = LU$ 分解，其中 L 为下三角矩阵。主元选取在计算中具有关键作用，部分主元法可有效限制舍入误差放大，使结果在浮点环境中保持稳定。

2.2 逆矩阵求解中的初等变换

对于可逆矩阵 A，构造增广矩阵 [A|I]，通过连续的行变换把左侧的 A 变为单位矩阵 I，右侧也就变成了 A^{-1}。表示为 $[E_kE_{k-1}\cdots E_1A \mid E_kE_{k-1}\cdots E_1I] = [I \mid A^{-1}]$。如果主元太小，计算就会不稳定，产生较大的舍入误差。对于病态矩阵，该过程会引起严重的数值漂移，所以常常要结合主元策略或者矩阵预处理来提高稳定性。

2.3 线性方程组求解中的应用

线性方程组 $AX=b$ 的求解可以看作是矩阵初等变换的直接应用，行变换可以把系统变为等价形式 $E_kE_{k-1}\cdots E_1AX=E_kE_{k-1}\cdots E_1b$，当 A 被化为阶梯形矩阵时就可以得到未知向量 X。当矩阵的秩小于未知数的个数时，系统有无穷多解；当矩阵的秩小于增广矩阵的秩时，系统无解。矩阵的初等变换给出系统化的求解方法，也是 QR 分解、LU 分解、

SVD 等现代算法的理论基础，使线性系统的数值解具有结构性和稳定性。

3 误差来源与传播机制

3.1 舍入与截断误差的分类研究

数值计算中误差不可避免，主要有舍入误差和截断误差。舍入误差是由于计算机对实数的有限表示造成的，所有的浮点数只能近似真实值，例如实数 0.1 的二进制展开为无限循环小数，计算机只能存储其近似值，从而引入微小误差。每次加减乘除运算都会产生新的舍入误差，多次迭代或者长链运算中这些误差不断累积，导致最终结果偏离理论值。双精度浮点数虽然可以提供大约 15 ~ 16 位有效数字，但是在高维矩阵计算中仍然难以避免数值漂移 [3]。

截断误差是由于算法的近似性造成的。连续计算被离散化时，高阶项被忽略或者步长过大，就会有截断偏差。截断误差的大小与算法阶数、收敛速度有关，主要出现在迭代求逆、特征分解等过程中。总体来说，舍入误差由计算环境决定，截断误差由算法结构决定，二者叠加影响矩阵初等变换的稳定性和精度 [4]。

3.2 初等变换的误差累积分析

矩阵初等变换由多步线性操作组成，每步浮点误差均可能在后续运算中被放大。设矩阵经第 k 步变换后为 A(k)，则浮点计算结果为 $\tilde{A}^{(k)} = A^{(k)} + \Delta A^{(k)}$。误差不单来自单步舍入，还会受到前期变换的影响而逐步累积。

在高斯消元法中，当主元过小时，除法运算放大误差，使舍入误差呈指数增长。为抑制这种现象，采取部分主元选取的办法，通过行交换保证主元是列中绝对值最大的那个，降低误差放大倍数。对于病态矩阵或者高维稀疏矩阵，即使采用主元策略，误差传播也比较大。如果累积误差不加控制，会破坏矩阵的对称性或者正定性，导致算法的数值特性不稳定。经验表明，当误差增长不超过 $O(n^2\varepsilon)$ 时，算法可以认为是稳定的。

3.3 误差传播模型与定量评估

误差传播模型用于刻画误差在计算中的传递规律，可形式化表示为 $\Delta x = P\Delta A + Q\Delta b + R\varepsilon$。其中 P、Q、R 为灵敏度矩阵，表示系统对各个误差源的响应强度。矩阵初等变换中误差传播常常具有非线性特点，尤其当主元极小或者矩阵病态时，误差会急剧增大，出现灾难性抵消现象，使得有效数字急剧减少。误差评价一般结合条件数 $\kappa(A)$ 和后向误差 β 来判断。当 β 与机器精度 ε 同数量级时，算法可以认为是稳定的；如果 β 明显变大，说明算法存在不稳定性。后向稳定性已经成为现代数值算法评价的重要标准，可以结合条件数、范数增长、残差等量化手段来综合判断算法的稳定性。

3.4 算法精度控制策略

矩阵初等变换中的误差无法消除，但可以通过优化策略有效抑制。主元调整机制可以控制误差传播，全主元法最稳定但是计算量较大；数值缩放、归一化能平衡矩阵量级，防止中间结果溢出或者下溢；精度增强、残差修正通过高精度计算、误差迭代提高计算结果的准确度；对于病态问题，可以采用再正则化技术改善条件数，抑制误差爆发 [5]。

4 数值实验与结果分析

4.1 实验设计与测试矩阵

为了验证矩阵初等变换法在计算中的稳定性以及误差控制性能，设计了多组 MATLAB 数值实验。测试矩阵包括三类：随机矩阵 U(0，1)、Hilbert 病态矩阵、对角占优矩阵，矩阵规模为 n=10、50、100。全部实验都在双精度环境下进行，重复运行 50 次以减小随机误差，输入向量 b 为全 1 列向量，运行平台为 MATLABR2023b，处理器 i7-12700H，内存 16GB，保证实验的科学性与重复性。

4.2 稳定性测试结果分析

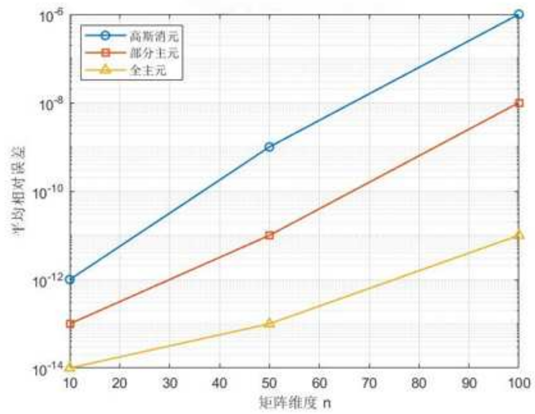

图 1 不同算法在各类矩阵上的稳定性测试结果曲线

实验结果表明，高斯消元法在随机矩阵上表现稳定，但在 Hilbert 矩阵中误差急剧上升；部分主元法改善了稳定性，全主元法在所有测试中误差保持在机器精度水平。图 1 展示各种算法的平均相对误差曲线，纵轴为 $\log_{10}$ 误差，横轴为矩阵维度。可以看出高斯法的误差随着维度线性增加，全主元法的误差几乎不变。随着条件数的增大，高斯法的误

差达到 10^{-3}，全主元法的误差仍小于 10^{-11}，虽然运算时间增加约 15%，但是稳定性提高了很多。

实验结果进一步显示，随着矩阵条件数增大，误差增长的速度加快。对于 Hilbert 矩阵，高斯消元法的误差甚至达到 10^{-3} 量级，而全主元法仍能维持在 10^{-11} 以下。这说明在浮点运算环境中，算法的鲁棒性与主元策略密切相关。从时间复杂度角度看，全主元法较慢约 15%，但在稳定性提升方面具有明显优势。

4.3 误差行为的对比与讨论

为了比较算法的误差规律，统计部分主元法和全主元法的平均误差，结果如表 5.1 所示。全主元法的误差波动最小、方差最低，说明它的误差传播控制更好。

表 5.1 不同算法在各矩阵维度下的误差对比结果

矩阵规模 n	部分主元法误差	全主元法误差	误差方差比 (%)
10	1.2×10^{-12}	9.5×10^{-14}	5.1
50	4.6×10^{-10}	2.1×10^{-12}	6.4
100	7.8×10^{-8}	1.9×10^{-11}	8.2

随着维度的增加，部分主元误差增长到全主元的上百倍。对于对角占优矩阵，二者相差不大，说明当矩阵条件良好时算法差异的影响较小。全主元法在复杂问题中精度最高，稳定性最好。

结论

本文系统地研究了矩阵初等变换法的稳定性以及误差特性。从理论和实验两方面可以得出，主元策略、条件数、数值尺度都会对算法的稳定性产生较大影响。高斯法计算效率高但是稳定性较差，部分主元法可以明显改善，全主元法既具有高精度又具备鲁棒性。研究给高可靠数值算法设计提供理论和实践依据，对科学计算、工程仿真等有重要参考价值。

参考文献

[1] 李敏丽 . 矩阵初等变换方法在高等代数中的应用 [J]. 数学学习与研究 ,2021,(15):6–7.

[2] 冯依虎 , 杨星星 . 逆矩阵若干求解方法的类比探究 [J]. 赤峰学院学报 (自然科学版),2021,37(05):1–5.

[3] 王翠翠 . 究 [J]. 教育教学论坛 ,2020,(47):284–286.

[4] 张芳英 , 朱睦正 . 矩阵初等变换在高等代数中的应用及教学研究 [J]. 河西学院学报 ,2020,36(02):117–122.

[5] 张楠 , 梅月兰 , 王双 . 实系数多项式因式分解的一种矩阵初等变换法 [J]. 湖北理工学院学报 ,2020,36(01):62–64.