1000-9825/2004/15(05)0650 ©2004 Journal of Software 软 件 学 报 Vol.15, No.5

矩阵条件数及高斯算法平滑分析的进一步研究*

杨智应+, 朱 洪, 宋建涛

(复旦大学 计算机科学与工程系,上海 200433)
(复旦大学 智能信息处理实验室,上海 200433)

Further Study on the Smoothed Analysis of Condition Number of Matrix and Gaussian Algorithm

YANG Zhi-Ying+, ZHU Hong, SONG Jian-Tao

(Department of Computer Science and Engineering, Fudan University, Shanghai 200433, China)

(Intelligent Information Processing Laboratory, Fudan University, Shanghai 200433, China)

+ Corresponding author: E-mail: 011021391@fudan.edu.cn, zhiy_yang@163.net, http://www.fudan.edu.cn

Received 2003-03-24; Accepted 2003-10-08

Yang ZY, Zhu H, Song JT. Further study on the smoothed analysis of condition number of matrix and Gaussian algorithm. *Journal of Software*, 2004,15(5):650~659.

http://www.jos.org.cn/1000-9825/15/650.htm

Abstract: Smoothed analysis aims to explain the success of algorithms that work well in practice while performing poorly in worst-case analysis. High performance computers have been used extensively in solving large scale linear systems and decomposition of matrix. The simplest and most implemented method of solving linear systems is Gaussian elimination (Gaussian algorithm). Natural implementations of Gaussian elimination use $O(n^3)$ arithmetic operations to solve a system of n linear equations in n variables. If the coefficients of these equations are specified using m bits, in the worst case it suffices to perform the elimination using mn bits of precision, because the elimination may produce large intermediate entries. However, a large number of experiments in numerical analysis have indicated that this high precision is necessary rarely in practice. Huge condition number and growth factors of matrix are always the main roots that make the matrix ill-conditioned and lead to produce a large error in solution. Let A be the matrix of independent Gaussian random variables centered at $\overline{A}$, each of the variances $\sigma^2 \le 1$, $K(A)$ be the condition number of matrix A. Sankar et al. performed a smoothed analysis of condition numbers and growth factors of matrices. They showed that for any matrix $\overline{A} \in R^{n\times n}$ and $\alpha>0$, $\Pr[K(A)\ge\alpha] \le \frac{3.64n\left(1+4\sqrt{\log(\alpha)}\right)}{\alpha\sigma}$, and the smoothed complexity of bits of precision needed to solve $Ax=b$ to m bits of accuracy using Gaussian algorithm is at most: $m+7\log_2(n)+3\log_2(1/\sigma)+\log_2\log_2 n+7$. They made some mistake in their proof. The mistake is corrected in this paper, and the smoothed complexity of bits of precision is corrected as: $m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367$. Two inequalities on norm of the matrix and product of the two random variables respectively are presented, and the smoothed analysis of condition number of matrix is

* Supported by the National Natural Science Foundation of China under Grant No.60273045 (国家自然科学基金); the Shanghai Science and Technology Development Fund of China under Grant No.025115032 (上海市科技发展基金)

作者简介: 杨智应(1964－),男,广西钟山人,博士,副教授,主要研究领域为算法设计与分析;朱洪(1939－),男,教授,博士生导师,主要研究领域为算法与复杂性,量子计算,计算密码学;宋建涛(1976－),男,博士,主要研究领域为 P2P 计算.

simplified to $\Pr[K(A)\ge\alpha]\le\frac{6\sqrt{2}n^2}{\alpha\cdot\sigma}$. The conjecture $\Pr[K(A)\ge\alpha]\le O\left(\frac{n}{\alpha\sigma}\right)$ is solved to some extent. The smoothed complexity of bits of precision is improved to $m+8\log_2 n+3\log_2(1/\sigma)+7$. And the experimental results show that the new smoothed analysis results are much better.
Key words: smoothed complexity; condition number of matrix; growth factor of matrix; smoothed precision

摘 要: 算法的复杂度平滑分析是对许多算法在实际应用中很有效但其最坏情况复杂度却很糟这一矛盾给出的更合理的解释.高性能计算机被广泛用于求解大规模线性系统及大规模矩阵的分解.求解线性系统的最简单且容易实现的算法是高斯消元算法(高斯算法).用高斯算法求解 n 个方程 n 个变量的线性系统所需要的算术运算次数为 $O(n^3)$.如果这些方程中的系数用 m 位表示,则最坏情况下需要机器位数 mn 位来运行高斯算法.这是因为在消元过程中可能产生异常大的中间项.但大量的数值实验表明,在实际应用中,需要如此高的精度是罕见的.异常大的矩阵条件数和增长因子是导致矩阵 $\overline{A}$ 病态,继而导致解的误差偏大的主要根源.设 $\overline{A}$ 为任意矩阵, A 是 $\overline{A}$ 受到微小幅度的高斯随机扰动所得到的随机矩阵,方差 $\sigma^2\le 1$ .Sankar 等人对矩阵 A 的条件数及增长因子进行平滑分析,证明了 $\Pr[K(A)\ge\alpha]\le\frac{3.64n\left(1+4\sqrt{\log(\alpha)}\right)}{\alpha\sigma}$.在此基础上证明了运行高斯算法输出具有 m 位精度的解所需机器位数的平滑复杂度为 $m+7\log_2(n)+3\log_2(1/\sigma)+\log_2\log_2 n+7$. 在上述结果的证明过程中存在错误,将其纠正后得到以下结果: $m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367$. 通过构造两个分别关于矩阵范数和随机变量乘积的不等式,将关于矩阵条件数的平滑分析结果简化到 $\Pr[K(A)\ge\alpha]\le\frac{6\sqrt{2}n^2}{\alpha\cdot\sigma}$.部分地解决了 Sankar 等人提出的猜想: $\Pr[K(A)\ge\alpha]\le O\left(\frac{n}{\alpha\cdot\sigma}\right)$.并将运行高斯算法输出具有 m 位精度的解所需机器位数的平滑复杂度降低到 $m+8\log_2 n+3\log_2(1/\sigma)+7$. 实验结果表明,所得到的平滑复杂度更好.
关键词: 平滑复杂度;矩阵条件数;矩阵增长因子;平滑精度
中图法分类号: TP301 **文献标识码**: A

算法的复杂度分析是软件系统设计与分析中的一个重要环节.为了分析算法的复杂度,我们常常对算法进行最坏情况复杂度分析.长期以来,人们一直为这样的现象所困惑,有很多算法按通常的(最坏情况)分析具有很坏的复杂度(通常是指数级的),但在实际应用中却很有效.这类算法的一个典型代表就是优化问题的单纯形算法.20 世纪 80 年代初,人们用随机方法分析了单纯形算法,并证明了对输入参数的一系列分布,单纯形算法的期望执行时间收敛于多项式复杂度.但是,在随后 20 年里,很多人的研究指出,这种统计平均的多项式复杂度并不能很好地解释单纯形算法的实际有效性,因为它的输入参数的统计分布具有很特殊的性质,在实际应用中很罕见.

另外,算法的平均复杂度与输入实例空间上的概率分布有关,不同的分布其相应的平均复杂度未必相同,而且输入实例空间上的概率分布一般不容易得知.为了更合理地解释这样的现象,美国麻省理工学院的 Daniel A. Spielman 与波士顿大学的滕尚华提出了计算机算法的平滑复杂度(smoothed complexity)概念.Spielman 和 Teng 在 STOC 2001 的论文[1]中提出了一种新的算法复杂度分析方法,即算法的平滑复杂度分析.他们的思路是这样的,在分析一个算法时,对输入实例加上合理的微小幅度的随机“扰动”,然后分析执行时间与输入规模及扰动的关系,由此得出的时间复杂度称为该算法的时间平滑复杂度.分析算法的平滑复杂度,其目的在于探测输入实例空间中是否含有最坏输入实例的稠密区域.Spielman 和 Teng 在文献[1]中分析了“两阶段投影-顶点”单纯形算法的平滑复杂度,证明了“两阶段投影-顶点”单纯形算法在具有给定规模的所有可能的输入实例的局部随机邻域上,其最高平均时间复杂度是多项式的,更合理地解释了单纯形算法虽然具有很坏的最坏情况复杂度(时间复杂度为指数级),但在实际应用中却很有效之间的矛盾.自文献[1]发表以来,这一领域的研究在理论计算机界引起了极大的关注并取得了一系列的成果[2].

矩阵条件数和增长因子是高性能计算中求解大规模线性系统 $\overline{A}\boldsymbol{y}=\boldsymbol{b}$ 时需要考虑的重要因素.这是因为异

常大的条件数和增长因子是导致矩阵 $\overline{A}$ 病态继而导致求解误差偏大的主要根源.由于在实际的计算环境中存在着这样或那样的随机干扰,因此,人们自然而然地会提出这样的问题:在有微小随机干扰的情况下,矩阵是否会出现异常大的条件数和增长因子?除此之外,如果要求高斯消元法输出具有 m 位精度的解,平均需要多少位的机器精度?最近,Sankar,Spielman 和 Teng 在文献[3]中对矩阵条件数和增长因子进行平滑分析,证明了 A 不可能有很大的条件数及很大的增长因子.在此基础上对求解线性系统 $\overline{A}\boldsymbol{y}=\boldsymbol{b}$ 的高斯顺序消元法(简称高斯算法)所需机器精度位数进行平滑分析,证明了高斯顺序消元法输出具有 m 位精度的解所用机器精度位数的期望值(也就是关于所需机器精度位数的平滑复杂度)不会超过

$$m+7\log_2(n)+3\log_2(1/\sigma)+\log_2\log_2 n+7 .$$

我们发现在他们的证明过程中有错误,纠正其中的错误后得到如下结果:

$$m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367 .$$

Sankar,Spielman 和 Teng 在文献[3]中提出了关于矩阵条件数的平滑分析结果 $\Pr[K(A)\ge\alpha]\le \frac{3.646n\left(1+4\sqrt{\log(\alpha)}\right)}{\alpha\sigma}$ 的改进猜想: $\Pr[K(A)\ge\alpha]\le O\left(\frac{n}{\alpha\sigma}\right)$.我们在本文中通过分别构造两个关于矩阵范数及随机变量乘积的不等式,将原结果简化到 $\Pr[K(A)\ge\alpha]\le\frac{6\sqrt{2}n^2}{\alpha\cdot\sigma}$,并在此基础上,将原文中给出的高斯顺序消元法输出具有 m 位精度的解所需机器精度位数的平滑复杂度降低到 $m+8\log_2 n+3\log_2(1/\sigma)+7$.

与原结果进行比较,我们的结果明显要好.另外,我们结合目前所做的工作,对算法的平滑复杂度分析的研究提出了一些有待进一步探索的工作.

本文第 1 节叙述算法的平滑复杂度分析基本概念.第 2 节简要介绍矩阵条件数及增长因子有关概念及矩阵的高斯随机扰动模型.第 3 节列出本文引用到的引理.第 4 节介绍有关矩阵条件数及高斯算法平滑分析的相关工作.第 5 节给出本文的主要结果及其详细证明.第 6 节进行结果的比较.第 7 节总结全文并提出一些有待进一步探索的工作.

下面给出本文中所用到的一些符号说明.

$\boldsymbol{x},\boldsymbol{y},\boldsymbol{b}$	列向量或算法的输入实例;	Λ	一个算法;
σ^2	随机变量的方差;	X_n	规模为 n 的输入实例空间;
λ	向量的长度;	α,β	常数;
$\boldsymbol{g},\boldsymbol{u},\boldsymbol{v}$	随机向量;	ζ	随机变量;
$\boldsymbol{v}$	单位向量;	μ_n	规模为 n 的输入实例空间 X_n 上的概率分布;
m,n	整数;	$\overline{A}$	普通矩阵;
$K(A)$	矩阵 A 的条件数;	R^n	n 维实数空间;
$R^{m\times n}$	$m\times n$ 实矩阵空间;	A	对矩阵 $\overline{A}$ 进行随机扰动而产生的随机矩阵.

1 算法的平滑复杂度分析基本概念

给定一个算法 Λ,令 X_n 表示规模为 n 的所有可能输入实例所构成的输入实例空间.对于输入实例 $\boldsymbol{x}\in X_n$,令 $C_\Lambda(\boldsymbol{x})$ 表示算法 Λ 关于输入实例 $\boldsymbol{x}$ 的复杂度.输入实例的规模 n 通常由包含在输入实例中的变量个数及表示各变量取值范围所需二进制位数之和表示.

算法 Λ 的最坏情况复杂度用下面的等式表示: $\sup\limits_{\boldsymbol{x}\in X_n}(C_\Lambda(\boldsymbol{x}))=f(n)$.

算法 Λ 的平均情况复杂度用下面的等式表示: $\mathop{E}\limits_{\boldsymbol{x}\xleftarrow{\mu_n}X_n}(C_\Lambda(\boldsymbol{x}))=f(n)$,其中 μ_n 是定义在 X_n 上的概率分布.

算法 Λ 的平滑复杂度用下面的等式表示: $\sup\limits_{\boldsymbol{x}\in X_n}\mathop{E}\limits_{\boldsymbol{g}}\left[C_\Lambda\left(\boldsymbol{x}+\boldsymbol{g}\cdot\|\boldsymbol{x}\|\right)\right]=f(n,\sigma)$,其中 $\boldsymbol{g}$ 是一个由均值为 0,方差为 σ^2 的高斯分布随机变量所组成的随机向量,$\|\boldsymbol{x}\|$表示 $\boldsymbol{x}$ 的模.我们把 $\boldsymbol{x}+\boldsymbol{g}\cdot\|\boldsymbol{x}\|$看作以输入实例 $\boldsymbol{x}$ 为中心的一个随机邻域(记为 $U_\sigma(\boldsymbol{x})$).该邻域是输入实例 $\boldsymbol{x}$ 受到随机扰动产生的.由高斯分布的性质容易看出,邻域 $U_\sigma(\boldsymbol{x})$ 的大小主

要由σ^2和$\|\boldsymbol{x}\|$的大小决定.由定义可知,当$\sigma^2=0$ 时,$U_\sigma(\boldsymbol{x})$就是输入实例 $\boldsymbol{x}$,此时平滑复杂度就是最坏情况复杂度;当σ^2 取得充分大时,$U_\sigma(\boldsymbol{x})$就可以认为是输入实例空间 X_n,此时平滑复杂度就是平均情况复杂度;在现实环境中,通常存在一些微小的随机干扰或误差.因此,在平滑分析中,一般σ^2 都取得很小,表示对输入实例 $\boldsymbol{x}$ 的一个微小幅度的随机扰动.我们从上述定义可以看到,算法 Λ 的平滑复杂度就是算法 Λ 在 X_n 中各输入实例 $\boldsymbol{x}$ 的随机邻域$U_\sigma(\boldsymbol{x})$上的平均复杂度的最大值.因此,如果一个算法具有较低的平滑复杂度,就说明在输入实例空间中不可能存在有最坏情况输入实例的稠密区域,也就说明算法在实际应用中通常有良好的效果.算法的平滑复杂度分析的一般方法是:根据具体问题确定输入实例的随机扰动模型,对具有给定规模的所有输入实例作微小幅度的随机扰动(通常是加上一个具有很小方差的随机向量),求出算法在各个输入实例的随机邻域$U_\sigma(\boldsymbol{x})$上的平均复杂度,然后对这些平均复杂度求最大值就得到算法的平滑复杂度.

2 矩阵条件数及增长因子的基本概念

定义 1. 设向量$\boldsymbol{x}\in R^n$,记$\|\bullet\|$为R^n到R^1的一个映射,对任给$\boldsymbol{x},\boldsymbol{y}\in R^n$,若映射$\|\bullet\|$满足:

(1) $\|\boldsymbol{x}\|\ge 0$,等号当且仅当$\boldsymbol{x}=0$时成立;

(2) 对常数α,$\|\alpha\boldsymbol{x}\|=|\alpha|\|\boldsymbol{x}\|$;

(3) $\|\boldsymbol{x}+\boldsymbol{y}\|\le\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$,

则称$\|\boldsymbol{x}\|$为向量 $\boldsymbol{x}$ 的范数.

向量 $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)^T$ 常用的范数是 Holder 范数或简称 p-范数,定义为$\|\boldsymbol{x}\|_p=(|x_1|^p+\ldots+|x_n|^p)^{1/p}$.

特别地,我们有以下最常用的几种范数:

$$\|\boldsymbol{x}\|_1=|x_1|+|x_2|+\ldots+|x_n|,\ \|\boldsymbol{x}\|_2=\sqrt{\sum_{i=1}^n x_i^2}=\sqrt{\boldsymbol{x}^T\boldsymbol{x}},\ \|\boldsymbol{x}\|_\infty=\max_{i=1,\ldots,n}|x_i|.$$

有关向量范数的不等式:$\|\boldsymbol{x}\|_\infty\le\|\boldsymbol{x}\|_1\le n\cdot\|\boldsymbol{x}\|_\infty$, $\|\boldsymbol{x}\|_\infty\le\|\boldsymbol{x}\|_2\le\sqrt{n}\cdot\|\boldsymbol{x}\|_\infty$.

定义 2. 设$\overline{A}\in R^{n\times m}$($\overline{A}$为$n\times m$矩阵)对应一个非负实数$\|\overline{A}\|$,若满足下列条件:

(1) $\|\overline{A}\|\ge 0$,当且仅当$\overline{A}=0$时等号成立;

(2) 对任意常数α,$\|\alpha\overline{A}\|=|\alpha|\cdot\|\overline{A}\|$;

(3) 成立三角不等式$\|\overline{A}+\overline{B}\|\le\|\overline{A}\|+\|\overline{B}\|$ ($\overline{A},\overline{B}\in R^{n\times m}$),

则称$\|\overline{A}\|$为矩阵 $\overline{A}$ 的范数.

对于矩阵 $\overline{A}$,我们定义$\|\overline{A}\|_2=\sup\limits_{x\ne 0}\dfrac{\|\overline{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$, $\|\overline{A}\|_\infty=\sup\limits_{x\ne 0}\dfrac{\|\overline{A}\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty}$, $\|\overline{A}\|_1=\sup\limits_{x\ne 0}\dfrac{\|\overline{A}\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_1}$.

下面给出几个常用的范数表达式:

(1) $\|\overline{A}\|_1=\max\limits_{j=1,2,\ldots,n}\sum\limits_{i=1}^n|\overline{a}_{ij}|$;

(2) $\|\overline{A}\|_2=\sqrt{\lambda_{\max}(\overline{A'}\overline{A})}$;

(3) $\|\overline{A}\|_\infty=\max\limits_{1\le i\le n}\sum\limits_{j=1}^m\|\overline{a}_{ij}\|$,

其中 $\lambda_{\max}(\overline{A'A})$ 表示 $\overline{A'A}$ 的最大特征值.

有关矩阵范数的不等式:

(1) $\dfrac{1}{\sqrt{n}}\|A\|_2\le\|A\|_\infty\le\sqrt{n}\|A\|_2$;

(2) $\dfrac{1}{n}\|A\|_\infty\le\|A\|_1\le n\cdot\|A\|_\infty$, $A\in R^{n\times m}$.

定义 3. 设 $\overline{A}$ 是一个 $n\times n$ 矩阵, $K(\overline{A})=\left\|\overline{A}^{-1}\right\|\cdot\left\|\overline{A}\right\|$ 称为矩阵 $\overline{A}$ 的条件数.

定义 4. 设 $\overline{A}$ 是一个 $n\times n$ 矩阵, $\overline{A}$ 的高斯随机扰动记为 A,是指对 $\overline{A}$ 的每一个元素加上相互独立、均值为 0、方差为 σ^2 的高斯随机变量而得到的随机矩阵.

因为$|A|$为连续型随机变量,因此 A 为奇异矩阵($|A|=0$)的概率为 0.在以下的讨论中,我们总是假定$|A|\neq 0$.

定义 5. 设$\overline{A}=\overline{L}\cdot\overline{U}$是$\overline{A}$的L-U分解,则称$\rho_{\overline{U}}(\overline{A})=\left\|\overline{U}\right\|_\infty\Big/\left\|\overline{A}\right\|_\infty$, $\rho_{\overline{L}}(\overline{A})=\left\|\overline{L}\right\|_\infty$为矩阵$\overline{A}$的增长因子.

定义 6[4,5]. 设 A 为 $n\times n$ 矩阵,如果在方程 $A\boldsymbol{x}=\boldsymbol{b}$ 中,矩阵 A 和右端项 $\boldsymbol{b}$ 的微小变化引起解向量 $\boldsymbol{x}$ 的变化 $\|\delta\boldsymbol{x}\|$ 很大,则称矩阵 A 为关于解方程组和矩阵求逆的病态矩阵;称相应的方程组为病态方程组.

例如,对于方程组 $A\boldsymbol{x}=\boldsymbol{b}$,如果仅考虑对右端项 $\boldsymbol{b}$ 的扰动对方程组的解的影响,设其扰动方程组为 $A(\boldsymbol{x}+\delta\boldsymbol{x})=\boldsymbol{b}+\delta\boldsymbol{b}$,则 $\delta\boldsymbol{x}=A^{-1}\delta\boldsymbol{b}$. 由此推得 $\dfrac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\le\left\|A^{-1}\right\|\cdot\|A\|\dfrac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}=K(A)\cdot\dfrac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}$.可见, $K(A)$ 越大,方程组右端项变化引起的解向量相对误差就可能越大.即 $K(A)$ 越大, A 对求解方程组来说就越可能呈病态.由此可见,矩阵条件数是判别一个矩阵是否病态的重要依据.

例如,给定方程组 $\begin{pmatrix}1 & 2 & 9\\ 3000 & 2000 & 1000\\ 4/10^6 & 3/10^6 & 2/10^6\end{pmatrix}\cdot\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}1\\ 2000\\ 3/10^6\end{pmatrix}$,用 MATLAB 函数计算系数矩阵的条件数和求解方程组得到 $K(A)=1.9201\times10^{10}$, $\boldsymbol{x}=(-0.16667,1.3333,-0.16667)^T$. 由于 A 的条件数非常大,因此方程组变态的可能性非常大.将 a_{33} 改为 $3/10^6$, b_3 改为 $4/10^6$,扰动量级均为 $1/10^6$.扰动得到方程组的解为 $\boldsymbol{x}=(-9.5000,16.5000,-2.5000)^T$. 显然,该方程组为病态方程组.

约定. 在下面的讨论中,范数 $\|\bullet\|$ 表示 $\|\bullet\|_2$.

3 本文用到的几个引理

引理 1. 令 $(a_1,a_2,...,a_n)$ 是从 R^n 中均匀随机抽取的一个单位向量,则对于 $c\le1$, $\Pr[a_1^2\ge c/n]\ge\Pr[\zeta^2>c]$,其中 ζ 是一个均值为 0,方差为 1 的高斯随机变量.

引理 2. 令 ζ 是一个正随机变量,若存在 $\alpha\ge1$,对任意给定的正数 β , $\Pr_A[\zeta\ge\beta]\le\dfrac{\alpha}{\beta}$,则 $E_A[\log\zeta]\le1+\log\alpha$.

以上各引理的证明在文献[3]中已给出,在此从略.

4 矩阵条件数与增长因子平滑分析相关工作

Sankar,Spielman 和 Teng 在文献[3]中对矩阵 A 的条件数与增长因子进行平滑分析,他们得到的主要结果有:

定理 1. 令 $\overline{A}$ 为 $n\times n$ 矩阵, A 是由相互独立且方差为 σ^2 的高斯随机变量所组成的随机矩阵,均值为 $\overline{A}$, $\boldsymbol{v}$ 是任意的单位向量,则对于任意给定正数 α , $\Pr\left[\left\|A^{-1}\boldsymbol{v}\right\|>\alpha\right]<\sqrt{\dfrac{2}{\pi}}\dfrac{1}{\sigma\alpha}$, $\Pr\left[\left\|A^{-1}\right\|>\alpha\right]<\dfrac{1.823\sqrt{n}}{\alpha\sigma}$.

定理 2. 令 $\overline{A}$ 为 $n\times n$ 矩阵,且$\left|\overline{a}_{ij}\right|_{\max}\le1$. A 是由相互独立且方差为 $\sigma^2\le1$ 的高斯随机变量组成的随机矩阵,均值为 $\overline{A}$,则对于任意给定正数 α,

(1) $\Pr[K(A)>\alpha]<\dfrac{3.646n\left(1+4\sqrt{\log(\alpha)}\right)}{\alpha\sigma}$;

(2) $\Pr[\rho_U(A)>1+\alpha]<\dfrac{1}{\sqrt{2\pi}}\dfrac{n(n+1)}{\alpha\sigma}$;

(3) $\Pr[\rho_L(A)>\alpha]<\sqrt{\dfrac{3+\sqrt{2}}{\pi}\dfrac{n^3}{\alpha\sigma}}$.

定理 3. 令 $\overline{A}$ 为 $n\times n$ 矩阵,且$\left|\overline{a}_{ij}\right|_{\max}\le1$. A 是由相互独立且方差为 $\sigma^2\le1$ 的高斯随机变量组成的随机矩阵,

均值为 $\overline{A}$,则高斯顺序消元法输出 m 位精度解所需机器位数的平滑复杂度至多为

$$m+7\log_2(n)+3\log_2(1/\sigma)+\log_2\log_2 n+7.$$

5 本文的主要结果及其证明

引理 3. 设 ζ 是一个正随机变量,若存在正数 α,β ,且 $\alpha\ge 1$,使得对于任意给定的正数 γ ,有以下等式成立: $\Pr[\zeta\ge\gamma]\le\dfrac{\alpha\left(1+\beta\sqrt{\log\gamma}\right)}{\gamma}$,则 $E[\log\zeta]\le\log\alpha+1+\beta\sqrt{\log\alpha}+\dfrac{\beta}{2\sqrt{\log\alpha}}$.

引理 3 是文献[3]中的引理 5.3 的纠正.文献[3]中引理 5.3 的结果为

$$E[\log\zeta]\le\log\alpha+\frac{\beta}{\sqrt{2}}\left(\frac{2\log\alpha}{\alpha}+\frac{1}{\alpha\sqrt{\log\alpha}}\right)+1.$$

证明: $E[\log\zeta]=\int_0^\infty\Pr[\log\zeta\ge y]\mathrm{d}y$

$$=\int_0^\infty\min\left(1,\frac{\alpha(1+\beta\sqrt{y})}{\mathrm{e}^y}\right)\mathrm{d}y$$

$$\le\int_0^{\log\alpha}\mathrm{d}y+\int_{\log\alpha}^\infty\frac{\alpha(1+\beta\sqrt{y})}{\mathrm{e}^y}\mathrm{d}y$$

$$\le\log\alpha+1+\alpha\int_{\log\alpha}^\infty\frac{\beta\sqrt{y}}{\mathrm{e}^y}\mathrm{d}y.$$

注:文献[3]中因漏掉了 $\int_{\log\alpha}^\infty\frac{\beta\sqrt{y}}{\mathrm{e}^y}\mathrm{d}y$ 前面的系数 α 而导致错误的结果.

令 $y=\dfrac{z^2}{2}$,

$$\int_{\log\alpha}^\infty\frac{\beta\sqrt{y}}{\mathrm{e}^y}\mathrm{d}y=\frac{\beta}{\sqrt{2}}\int_{\sqrt{2\log\alpha}}^\infty z^2\cdot\mathrm{e}^{-z^2/2}\mathrm{d}z$$

$$=\frac{\beta}{\sqrt{2}}\left(\frac{\sqrt{2\log\alpha}}{\alpha}+\int_{\sqrt{2\log\alpha}}^\infty\mathrm{e}^{-z^2/2}\mathrm{d}z\right)$$

$$\le\frac{\beta}{\sqrt{2}}\left(\frac{\sqrt{2\log\alpha}}{\alpha}+\frac{1}{\alpha\sqrt{2\log\alpha}}\right)\text{(利用}\int_y^\infty\mathrm{e}^{-z^2/2}\mathrm{d}z\le\frac{\mathrm{e}^{-y^2/2}}{y}\text{)},$$

即
$$E[\log\zeta]\le\log\alpha+1+\beta\sqrt{\log\alpha}+\frac{\beta}{2\sqrt{\log\alpha}}.$$
□

引理 4(更正后的高斯顺序消元法精度的平滑复杂度). 令 $\overline{A}$ 为 $n\times n$ 矩阵,且 $|\overline{a}_{ij}|_{\max}\le 1$. A 是由相互独立且方差为 $\sigma^2\le 1$ 的高斯随机变量所组成的随机矩阵,均值为 $\overline{A}$,则高斯顺序消元法输出 m 位精度解所需机器位数的平滑复杂度至多为 $m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367$.

证明:高斯消元法输出 m 位精度解所需机器精度 ε 满足以下条件[1]: $5\cdot 2^m\cdot n\cdot\rho_L(A)\cdot\rho_U(A)\cdot K(A)\cdot\varepsilon\le 1$.

由此得: $2.33+m+\log_2 n+\log_2(\rho_L(A))+\log_2(\rho_U(A))+\log_2(K(A))\le\log_2(1/\varepsilon)$.

由引理 2 及定理 2 得: $E[\log_2(\rho_U(A))]\le 2\log_2 n+\log_2(1/\sigma)-0.325$ (原文中的结果);

由引理 2 及定理 2 得: $E[\log_2(\rho_L(A))]\le 3\log_2 n+\log_2(1/\sigma)+1.246$ (原文中的结果);

由引理 3 及定理 2 得: $E[\log_2(K(A))]\le\log_2 n+\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+4.116$.

于是,高斯顺序消元法输出 m 位精度解所需机器位数的平滑复杂度至多为

$$m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367. \qquad \square$$

定理 4. 设 $\boldsymbol{u}\in R^m,\boldsymbol{v}\in R^n$ 是相互独立的随机向量, $X(\boldsymbol{u},\boldsymbol{v}),Y(\boldsymbol{u})$ 是正随机变量,且有正数 α ,使得对于任意给定的正数 β ,有 $\Pr[X(\boldsymbol{u},\boldsymbol{v})>\beta]\le\frac{\alpha}{\beta}$,则 $\Pr[X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta]\le\frac{\alpha}{\beta}EY(\boldsymbol{u})$.

证明:我们对 $\boldsymbol{u}$ 为离散型随机向量给出证明,对于连续型的情况类似可证.此时, $Y(\boldsymbol{u})$ 为离散型随机变量,设其取值 $y_1,y_2,\ldots$,对于任意给定的正数 β ,

$$\Pr[X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta]=\sum_i\Pr[\{X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta\}\,|\,\{Y(\boldsymbol{u})=y_i\}]\cdot\Pr[Y(\boldsymbol{u})=y_i].$$

$$\begin{aligned}
\Pr[\{X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta\}\,|\,\{Y(\boldsymbol{u})=y_i\}]&=\frac{\Pr[\{X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta\}\cdot\{Y(\boldsymbol{u})=y_i\}]}{\Pr[Y(\boldsymbol{u})=y_i]}\\
&=\frac{\sum_{\boldsymbol{w}\in\{z|Y(z)=y_i\}}\Pr[\{X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta\}\cdot\{\boldsymbol{u}=\boldsymbol{w}\}]}{\Pr[Y(\boldsymbol{u})=y_i]}\\
&=\frac{\sum_{\boldsymbol{w}\in\{z|Y(z)=y_i\}}\Pr\left[\left\{X(\boldsymbol{w},\boldsymbol{v})>\frac{\beta}{y_i}\right\}\cdot\{\boldsymbol{u}=\boldsymbol{w}\}\right]}{\Pr[Y(\boldsymbol{u})=y_i]}\\
&=\frac{\sum_{\boldsymbol{w}\in\{z|Y(z)=y_i\}}\Pr\left[X(\boldsymbol{w},\boldsymbol{v})>\frac{\beta}{y_i}\right]\cdot\Pr[\boldsymbol{u}=\boldsymbol{w}]}{\Pr[Y(\boldsymbol{u})=y_i]}\ (\text{由文献[6]及}\boldsymbol{w},\boldsymbol{v}\text{相互独立})\\
&\le\frac{\sum_{\boldsymbol{w}\in\{z|Y(z)=y_i\}}\Pr\left[X(\boldsymbol{u},\boldsymbol{v})>\frac{\beta}{y_i}\right]\cdot\Pr[\boldsymbol{w}=\boldsymbol{w}]}{\Pr[Y(\boldsymbol{u})=y_i]}\ \left(\boldsymbol{w}\text{取定},\left\{X(\boldsymbol{w},\boldsymbol{v})>\frac{\beta}{y_i}\right\}\subseteq\left\{X(\boldsymbol{u},\boldsymbol{v})>\frac{\beta}{y_i}\right\}\right)\\
&\le\frac{\sum_{\boldsymbol{w}\in\{z|Y(z)=y_i\}}\frac{\alpha\cdot y_i}{\beta}\cdot\Pr[\boldsymbol{u}=\boldsymbol{w}]}{\Pr[Y(\boldsymbol{u})=y_i]}\\
&=\frac{\alpha\cdot y_i}{\beta}.
\end{aligned}$$

于是,
$$\Pr[X(\boldsymbol{u},\boldsymbol{v})\cdot Y(\boldsymbol{u})>\beta]\le\sum_i\frac{\alpha\cdot y_i}{\beta}\cdot\Pr[Y(\boldsymbol{u})=y_i]=\frac{\alpha}{\beta}\cdot EY(\boldsymbol{u}). \qquad \square$$

定理 5. 设 A 为任意的 $n\times n$ 矩阵, $\boldsymbol{u},\boldsymbol{v}$ 是单位向量,且 $\|A\boldsymbol{u}\|=\|A\|$, 则 $\|A\boldsymbol{v}\|\ge\|A\|\cdot|\langle\boldsymbol{u},\boldsymbol{v}\rangle|$.

证明:当 A 为零矩阵时,结论是显然的.设 A 为非零矩阵,则 A^TA 为半正定对称矩阵.因此, A^TA 的特征根 λ_i 均为非负实数($i=1,2,\ldots,n$).不妨设 λ_1 为 A^TA 的最大特征根.于是, $\lambda_1>0$.设 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\ldots,\boldsymbol{\eta}_n$ 为 A^TA 的分别属于 $\lambda_1,\lambda_2,\ldots,\lambda_n$ 且彼此正交的标准化特征向量,即 $\langle\boldsymbol{\eta}_i,\boldsymbol{\eta}_i\rangle=1,\langle\boldsymbol{\eta}_i,\boldsymbol{\eta}_j\rangle=0,i\ne j,i,j=1,2,\ldots,n.$

现以 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\ldots,\boldsymbol{\eta}_n$ 为基底,令 $\boldsymbol{u}=u_1\boldsymbol{\eta}_1+u_2\boldsymbol{\eta}_2+\ldots+u_n\boldsymbol{\eta}_n,\ \boldsymbol{v}=v_1\boldsymbol{\eta}_1+v_2\boldsymbol{\eta}_2+\ldots+v_n\boldsymbol{\eta}_n.$

由于 $\boldsymbol{u},\boldsymbol{v}$ 为单位向量,即 $\langle\boldsymbol{u},\boldsymbol{u}\rangle=\sum_{i=1}^n u_i^2=1,\langle\boldsymbol{v},\boldsymbol{v}\rangle=\sum_{i=1}^n v_i^2=1.$

$$(A^TA)\boldsymbol{u}=u_1\lambda_1\boldsymbol{\eta}_1+u_2\lambda_2\boldsymbol{\eta}_2+\ldots+u_n\lambda_n\boldsymbol{\eta}_n,\ (A^TA)\boldsymbol{v}=v_1\lambda_1\boldsymbol{\eta}_1+v_2\lambda_2\boldsymbol{\eta}_2+\ldots+v_n\lambda_n\boldsymbol{\eta}_n.$$

由矩阵范数的定义, $\|A\|=\sqrt{\lambda_1}$.

由假设,
$$\lambda_1=\|A\|^2=\|A\boldsymbol{u}\|^2=\langle A\boldsymbol{u},A\boldsymbol{u}\rangle=\langle\boldsymbol{u},A^TA\boldsymbol{u}\rangle=\sum_{i=1}^n\lambda_i u_i^2=\lambda_1\sum_{i=1}^n\frac{\lambda_i}{\lambda_1}u_i^2,$$

因此, $\sum_{i=1}^n\frac{\lambda_i}{\lambda_1}u_i^2=1$, 由 $\sum_{i=1}^n u_i^2=1,\ \lambda_i\le\lambda_1$ 可得: 或者 $u_i=0$,或者 $\lambda_i=\lambda_1$.

于是, $\boldsymbol{u}=(u_1,\sqrt{\frac{\lambda_2}{\lambda_1}}u_2,\ldots,\sqrt{\frac{\lambda_n}{\lambda_1}}u_n)$.由Cauchy-Schwarz不等式 得:

$$\|A\boldsymbol{v}\| = \sqrt{\langle A\boldsymbol{v}, A\boldsymbol{v}\rangle} = \sqrt{\langle \boldsymbol{v}, A^T A\boldsymbol{v}\rangle} = \sqrt{\sum_{i=1}^{n}\lambda_i v_i^2} = \sqrt{\lambda_1}\cdot\sqrt{\sum_{i=1}^{n}\frac{\lambda_i}{\lambda_1}v_i^2} \ge \sqrt{\lambda_1}\cdot\left|\sum_{i=1}^{n}\sqrt{\frac{\lambda_i}{\lambda_1}}u_i v_i\right|,$$

即

$$\|A\boldsymbol{v}\| \ge \sqrt{\lambda_1}\cdot|\langle \boldsymbol{u},\boldsymbol{v}\rangle| = \|A\|\cdot|\langle \boldsymbol{u},\boldsymbol{v}\rangle|.$$ □

定理 6(**矩阵条件数平滑分析结果的改简**). 令 $\overline{A}$ 为 $n\times n$ 矩阵,且 $|\overline{a}_{ij}|_{\max}\le 1$, A 是由相互独立且方差为 $\sigma^2\le 1$ 的高斯随机变量所组成的随机矩阵,均值为 $\overline{A}$,则对于给定的正数 α, $\Pr[K(A)\ge\alpha]\le\dfrac{6\sqrt{2}n^2}{\alpha\cdot\sigma}$.

证明:由定义可知,$\|A^{-1}\boldsymbol{x}\|$ 是定义在 n 维单位球面上的连续函数,因此存在单位向量 $\boldsymbol{u}$, s.t. $\|A^{-1}\boldsymbol{u}\| = \|A^{-1}\|$.这样的 $\boldsymbol{u}$ 以概率 1 惟一确定[1].令 $\boldsymbol{v}$ 是一个均匀随机抽取的单位向量,由定理 5 得到 $\|A^{-1}\boldsymbol{v}\|\ge\|A^{-1}\|\cdot|\langle \boldsymbol{u},\boldsymbol{v}\rangle|$.

于是

$$\|A\|\cdot\|A^{-1}\boldsymbol{v}\|\ge\|A^{-1}\|\cdot\|A\|\cdot|\langle \boldsymbol{u},\boldsymbol{v}\rangle| = K(A)\cdot|\langle \boldsymbol{u},\boldsymbol{v}\rangle|.$$

于是,对于任意给定的正数 α ,

$$\Pr\left[\|A^{-1}\boldsymbol{v}\|\cdot\|A\|\ge\alpha\sqrt{c/n}\right]\ge\Pr\left[K(A)\ge\alpha \text{ 且 } |\langle \boldsymbol{u},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\right]=\Pr[K(A)\ge\alpha]\cdot\Pr\left[|\langle \boldsymbol{u},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\,\big|\{K(A)\ge\alpha\}\right]$$

由于在 A 取定的条件下,满足 $\|A^{-1}\boldsymbol{u}\| = \|A^{-1}\|$ 的非零向量 $\boldsymbol{u}$ 以概率 1 惟一确定,并且在 $\boldsymbol{u}$ 固定的条件下,由引理 1,对于 $0<c\le 1$,

$$\Pr\left[|\langle \boldsymbol{u},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\right]\ge\Pr\left[|\zeta|\ge c\right] \tag{1}$$

其中,ζ为均值为 0,方差为σ^2 的高斯随机变量.

$$\begin{aligned}
\Pr\left[|\langle \boldsymbol{u},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\,\big|\{K(A)\ge\alpha\}\right] &= \frac{\Pr\left[\left\{|\langle \boldsymbol{u},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\right\}\cdot\{K(A)\ge\alpha\}\right]}{\Pr[K(A)\ge\alpha]}(\text{非零单位向位量}\boldsymbol{u}\text{满足}\|A^{-1}\boldsymbol{u}\|=\|A^{-1}\|)\\
&= \frac{\sum_{\overline{B}\in\{\overline{Z}|K(\overline{Z})\ge\alpha\}}\Pr\left[\left\{|\langle \boldsymbol{u}_{\overline{B}},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\right\}\cdot\left\{A=\overline{B}\right\}\right]}{\Pr[K(A)\ge\alpha]}(\text{其中},\|\overline{B}^{-1}\boldsymbol{u}_{\overline{B}}\|=\|\overline{B}^{-1}\|)\\
&= \frac{\sum_{\overline{B}\in\{\overline{Z}|K(\overline{Z})\ge\alpha\}}\Pr\left[\left\{|\langle \boldsymbol{u}_{\overline{B}},\boldsymbol{v}\rangle|\ge\sqrt{c/n}\right\}\right]\cdot \mathrm{P}\left[\left\{A=\overline{B}\right\}\right]}{\Pr[K(A)\ge\alpha]}(\overline{B}\text{取定},\boldsymbol{u}_{\overline{B}}\text{随之取定.而}\boldsymbol{v}\text{与}A\text{相互独立})\\
&\ge \frac{\sum_{\overline{B}\in\{\overline{Z}|K(\overline{Z})\ge\alpha\}}\Pr[|\zeta|\ge c]\cdot P\left[\left\{A=\overline{B}\right\}\right]}{\Pr[K(A)\ge\alpha]}(\text{根据式(1)推出})\\
&= \Pr[|\zeta|\ge c],
\end{aligned}$$

$$\Pr\left[\|A^{-1}\boldsymbol{v}\|\cdot\|A\|\ge\alpha\sqrt{c/n}\right]\ge\Pr[K(A)\ge\alpha]\cdot\Pr[|\zeta|\ge c],$$

因此,

$$\Pr[K(A)\ge\alpha]\le\frac{\Pr\left[\|A^{-1}\boldsymbol{v}\|\cdot\|A\|\ge\alpha\sqrt{c/n}\right]}{\Pr\left[|\zeta|\ge c\right]}.$$

把定理 4 中的向量 $\boldsymbol{u}$ 取为由 A 中所有元素构成的随机向量,定理 4 中的正随机变量 $X(\boldsymbol{u},\boldsymbol{v})$ 及 $Y(\boldsymbol{u})$ 分别取为 $\|A^{-1}\boldsymbol{v}\|$ 和 $\|A\|$. 由定理 1, $\Pr\left[\|A^{-1}\|>\alpha\right]<\dfrac{2\sqrt{n}}{\alpha\sigma}$. 由此可得 $\Pr\left[\|A^{-1}\boldsymbol{v}\|>\alpha\right]<\dfrac{2\sqrt{n}}{\alpha\sigma}$.

令$\|A\|=\|A\boldsymbol{w}\|$,其中$\|\boldsymbol{w}\|=1$,显然$\|A\|^2\le\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2$,由于$E\|A\|\le\sqrt{E\|A\|^2}$,$Ea_{ij}^2\le 1+\sigma^2$.而$\sigma^2<1$,因此,$E\|A\|\le\sqrt{2}n$.

于是,由定理 4 可得 $\Pr\left[\|A^{-1}\boldsymbol{v}\|\cdot\|A\|\ge\alpha\sqrt{c/n}\right]\le\dfrac{2\sqrt{2}n^2}{\alpha\cdot\sqrt{c}\cdot\sigma}$. 因此, $\Pr[K(A)\ge\alpha]\le\dfrac{2\sqrt{2}n^2}{\alpha\cdot\sqrt{c}\cdot\sigma\cdot\Pr\left[|\zeta|\ge c\right]}$.

由于 $\Pr[|\zeta|\ge 1]\ge\dfrac{1}{3}$,取 $c=1$,立即得到我们所要证明的结论. □

定理 7(**高斯消元法的平滑复杂度改进**). 令 $\overline{A}$ 为 $n\times n$ 矩阵,且 $|\overline{a}_{ij}|_{\max}\le 1$, A 是由相互独立且方差为 $\sigma^2\le 1$

的高斯随机变量所组成的随机矩阵,均值为 $\overline{A}$,则高斯顺序消元法输出 m 位精度的解所需机器位数的平滑复杂度至多为 $m+8\log_2 n+3\log_2(1/\sigma)+7$.

证明:高斯消元法输出 m 位精度结果所需机器精度 ε 满足如下条件[1]: $5\cdot 2^m\cdot n\cdot \rho_L(A)\cdot \rho_U(A)\cdot K(A)\cdot\varepsilon\le 1$.

于是有 $$2.33+m+\log_2 n+\log_2(\rho_L(A))+\log_2(\rho_U(A))+\log_2(K(A))\le\log_2(1/\varepsilon).$$

由引理 2 及定理 2 得到: $E[\log_2(\rho_U(A))]\le 2\log_2 n+\log_2(1/\sigma)-0.325$ (原文的结论)

由引理 2 及定理 2 得到: $E[\log_2(\rho_L(A))]\le 3\log_2 n+\log_2(1/\sigma)+1.246$ (原文的结论)

由引理 2 及定理 6 得到: $E[\log_2(K(A))]\le 2\log_2 n+\log_2(1/\sigma)+3.085$.

因此,高斯顺序消元法输出 m 位精度的解所需机器位数的平滑复杂度至多为 $m+8\log_2 n+3\log_2(1/\sigma)+7$. □

6 结果的比较

从定理 2 及定理 6 可以看到,改简前后关于矩阵条件数的平滑分析结果分别为

$$p_0=\Pr[K(A)\ge\alpha]\le\frac{3.646n(1+4\sqrt{\log(\alpha)})}{\alpha\sigma},\quad p_1=\Pr[K(A)\ge\alpha]\le\frac{6\sqrt{2}n^2}{\alpha\cdot\sigma}.$$

于是, $\frac{p_0}{p_1}=\frac{3.646(1+4\sqrt{\log(\alpha)})}{6\sqrt{2}\cdot n}$.对于给定的矩阵阶数 n ,当 $\alpha>2^{\left(\frac{6\sqrt{2}\cdot n-3.646}{14.584}\right)^2}$ 时,必有 $p_0>p_1$.其他情况有 $p_0\le p_1$.由此可见,对于给定的矩阵阶数 n, 当 α 取得足够大以后,我们的结果比原结果要好.在实际中,矩阵的阶数一般都较小.因此, α 越大,越能反映改简后结果的优越性能.下面,我们对 n 及 α 的一些取值情况进行对比,见表 1.

Table 1 Comparison of smoothed analysis of condition number of matrix between pre-simplification and post-simplification

表 1 改简前后矩阵条件数的平滑分析结果对比

n \ p_0/p_1 \ α	2^{100}	2^{150}	2^{200}	2^{250}	2^{300}	2^{400}
10	1.761 71	2.147 99	2.473 64	2.760 54	3.019 92	3.480 45
20	0.880 86	1.073 99	1.236 82	1.380 27	1.509 96	1.740 22
30	0.587 24	0.715 99	0.824 55	0.920 18	1.006 64	1.160 15
40	0.440 43	0.536 99	0.618 41	0.690 13	0.754 98	0.870 11

从引理 4 及定理 7 看到,改简前后关于高斯消元法输出具有 m 位精度的解所需的机器位数的平滑复杂度的界分别为

$$c_0=m+7\log_2 n+3\log_2(1/\sigma)+4\sqrt{2+\log_2 n+\log_2(1/\sigma)}+7.367,\ c_1=m+8\log_2 n+3\log_2(1/\sigma)+7.$$

于是, $c_0-c_1=4\sqrt{2+\log_2 n+\log_2(1/\sigma)}-\log_2(n)-1$.

下面我们对 n 及 σ 的一些取值情况进行比较,见表 2.

Table 2 Comparison of smoothed complexity of bits of precision needed to perform Gaussian algorithm between pre-simplification and post-simplification

表 2 改简前后所需的机器位数的平滑复杂度界的比较

n \ c_0-c_1 \ σ	0.01	0.001	0.000 1	0.000 01	0.000 001	0.000 000 1
50	8.475 780	10.141 68	11.656 56	13.055 28	14.361 07	15.590 299
100	7.995 946	9.611 701	11.088 60	12.457 29	13.738 54	14.947 240
200	7.499 359	9.069 253	10.510 91	11.851 41	13.109 46	14.298 629
300	7.201 726	8.746 544	10.168 72	11.493 52	12.738 57	13.916 743
400	6.987 542	8.515 281	9.924 105	11.238 10	12.474 16	13.644 717
500	6.819 753	8.334 628	9.733 351	11.039 14	12.268 37	13.433 115
600	6.681 616	8.186 220	9.576 850	10.876 05	12.099 78	13.259 839

上述实验结果表明,对于给定的矩阵阶数,随着方差取值的不断减小,两者的差距不断地增大.在平滑分析中,一般方差都取得非常小.

7 总结与展望

本文首先纠正文献[3]关于高斯顺序消元法输出具有 *m* 位精度的解所用机器精度位数的平滑复杂度证明过程中的错误,然后通过分别构造两个关于矩阵范数和随机变量乘积的不等式,改简原文中关于矩阵条件数平滑分析结果,进而改进原文关于高斯顺序消元法输出 *m* 位精度解所需机器精度位数的平滑复杂度,并将我们的结果与原结果进行比较.

算法的平滑复杂度分析是算法研究的一个新领域,下面我们提出一些有待进一步探索的工作:

(1) 从已有实验结果来看,有选主元操作的高斯消元法的精度比高斯顺序消元算法的精度要高.因此,从直觉上看,有选主元操作的高斯消元法输出 *m* 位精度解所需机器精度位数的平滑复杂度要比高斯顺序消元算法低得多,这是一项富于挑战性的研究工作.

(2) 除了单纯形算法以外,在很多优化问题(如二次规划问题、线性补问题 LP)、数值计算问题及人工智能领域中也常常碰到情况类似的算法,它们具有很坏的最坏情况复杂度(例如指数级的时间或空间复杂度),而实际应用中却很有效.需要对这样的算法作平滑复杂度分析.

(3) 在计算机资源管理中(如内存管理、作业调度等)存在很多算法(如在 Unix 及 Windows NT 中使用的 MLF 调度算法),这些算法在实际应用中表现出令人满意的性能,但它们的最坏情况性能分析却令人失望.因此,有必要对这些算法作平滑分析.

在实际的应用环境中,通常都存在这样或那样的随机干扰.我们认为,应用算法的平滑复杂度分析方法对求解实际问题的算法进行评价是一项非常有价值的研究工作,尤其是经济领域中所使用的算法.我们相信,随着对算法的平滑复杂度研究的不断深入,人们对算法的平滑复杂度分析的意义将有更好的认识.算法的平滑复杂度分析在系统开发与分析、经济领域及其他工程技术领域中将发挥其应有的作用.

References:

[1] Spielman DA, Teng SH. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. In: SIGACT, ed. Proc. of the 33rd Annual ACM Symp. on Theory of Computing. ACM Press, 2001. 296~305. http://www-math.mit.edu/~spielman/SmoothedAnalysis/papers.html

[2] Spielman DA, Teng SH. Smoothed analysis of algorithms. In: LI T, ed. Proc. of the 2002 Int'l Congress of Mathematicians, Abstracts of Plenary and Invited Lectures. Beijing: Higher Education Press, 2002. 144~145. http://www-math.mit.edu/~spielman/SmoothedAnalysis/papers.html

[3] Sankar A, Spielman DA, Teng SH. Smoothed analysis of the condition numbers and growth factors of matrices. In: Cucker F, Devore R, Olver P, eds. Proc. of the 2002 Conf. on the Foundations of Computational Mathematics. London: Cambridge University Press, 2002. http://www-math.mit.edu/~spielman/SmoothedAnalysis/papers.html

[4] Guan Z, Lu JF. Basic of Numerical Analysis. Beijing: Higher Education Press, 1998 (in Chinese).

[5] Shi MG, Gu LZ. Basic of Science and Engineering Computation. Beijing: Tsinghua University Press, 1999 (in Chinese).

[6] Fudan University. Basic of Probability Theory. Beijing: People's Education Press, 1979 (in Chinese).

附中文参考文献:

[4] 关治,陆金甫.数值分析基础.北京:高等教育出版社,1998.

[5] 施妙根,顾丽珍.科学与工程计算基础.北京:清华大学出版社,1999.

[6] 复旦大学.概率论.北京:人民教育出版社,1979.