线代数方程组的计算机解法

G. E. 福赛思　C. B. 莫勒 著

科学出版社

线代数方程组的计算机解法

G. E. 福赛思　C. B. 莫勒 著

徐树荣 译

科学出版社

1979

内 容 简 介

本书是介绍线代数方程组数值解法的专著. 全书共分二十五章，其中心为解紧凑存储系数矩阵的方程组的消去法——有迭代改善的高斯消去法，研究了它的误差，给出了用三种计算机语言编写的实用程序. 最后简略介绍了解非线性方程组的方法. 可供计算数学专业师生，工程技术人员及有关科学计算工作者参考.

G. E. Forsythe, C. B. Moler
COMPUTER SOLUTION OF LINEAR ALGEBRAIC SYSTEMS
Prentice-Hall, 1967

线代数方程组的计算机解法

G. E. 福赛思 C. B. 莫勒 著
徐树荣 译

*

科学出版社出版
北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1979 年 11 月第 一 版 开本：787×1092 1/32
1979 年 11 月第一次印刷 印张：4 3/4
印数：0001—20,100 字数：103,000

统一书号：13031・1112
本社书号：1564・13—1

定价： 0.50 元

译 者 序

本书叙述线代数方程组数值解法. 全书以求解紧凑存储系数矩阵的方程组的高斯消去法为中心，系统论述了有迭代改善的高斯消去法及其变形,研究了它的误差分析,给出了用三种语言编写的实用程序，最后简略叙述解线性和非线性方程组的其它方法.

本书的特点是主题突出，选材精练，与计算机联系紧密.反映了由电子计算机引起的现代计算数学发展的趋势.

本书作者之一福赛思（G. E. Forsythe）是数值分析方面的著名数学家. 另一位作者也是有相当经验的专家. 因此，尽管本书的篇幅不大，但却包含了丰富的内容. 是一本有特色的专著. 本书可作为计算数学和计算机科学系的教学参考书. 也可供从事数学应用的科学工作者和工程技术人员参考.

译者衷心感谢黄华同志给予的帮助和支持. 李荣华教授和林公豫同志审阅了本书全部译稿并对译文作了不少修改；薛彦才同志对部分译文提供了一些有益意见，译者在此向他们致以谢意. 本书译文主要根据英文原著，同时亦参考俄文译本. 由于译者水平所限，谬误之处难免，谨请读者指正.

徐 树 荣

于中国科学院沈阳计算所

原　　序

这本专著主要是作为学习数值分析和计算机科学的大学生的课本．它来源于斯坦福大学高年级数值分析课教材．这本教材几乎没有涉及到矩阵问题．作者建议读者利用计算中心的合适的程序，如果有疑问，就找有关专家答疑．与此相反，我们认为，为了使学生很好地使用这些程序并在需要的情况下加以修改，他们还需要了解矩阵程序．而且，后来我们的许多大学生自己改写或者翻译了计算中心的这些程序．为了弥补这一点，我们准备了一份关于矩阵计算的笔记，而这份笔记的部分内容构成了现在这本专著的基础．在第1章详细介绍了阅读本书必须具备的知识，它包括线性代数初步和程序设计的某些知识．

我们认为，除数值分析专家和计算机程序工作者外，大学生和从事数学规划、统计、工程以及涉及矩阵及其计算的很多其它领域的人都应该熟悉本书的大部分内容．

求解线性代数方程组是计算工作中经常遇到的．因为，线性函数研究得非常充分，而多变量的函数关系最通用的模型是线性函数，这就导致线性方程组．此外，求解非线性问题的大多数方法都归结为一系列的线性问题．由于线性系统如此之普遍，所以每一个计算实验室都应当有能力迅速而准确地求解线性代数方程组．毫不奇怪，这个要求并非都能满足．

如我们在第6章所指出的，线性方程组有多种类型．本书着重处理的一个重要类型是有紧凑存储的系数矩阵的方程

组．求解这个方程组的一个重要算法——有迭代改善的高斯消去法——的完整研究构成本书的中心内容．本书的最后几章简略叙述求解线性和非线性方程组的其它方法．

和其它的数学分支比较，数值分析有以下两个特点：

(1) 要考虑算法需要的机器时间和存储量这样一些经济指标．

(2) 要进行误差分析，误差是由计算机不同形式的有限精度运算引起的．

上述第一个考虑使高斯消去法成为求解紧凑线性方程组的最好方法．但是，还有消去法的各种变形，而误差分析用来指导选择这些变形中的一个．

除了用数值例子和误差分析说明高斯消去法外，还介绍了我们认为是最好的计算机算法．为了使算法得到广泛应用，我们给出了用三种语言，ALGOL 60，ASA FORTRAN 与 PL/I 编写的算法．在第 17 章，我们把注意力集中在翻译程序时由于语言或者机器结构不同而产生困难的那些地方．我们认为，这些程序以及伴随的讨论是本书的重要组成部分．

我们将本书献给威尔金森 (J. H. Wilkinson)，他在数值分析方面有很多重要贡献．专家们最好去研究他的两本关于矩阵计算的书．然而，对于低年级大学生来说，学习这些广博的材料是不容易的．本书的主要部分致力于阐述并推导某些算法和威尔金森提出的分析方法．

G. E. 福赛思

C. B. 莫　勒

目　　录

1. 读者具备的基础知识和本书的目的

我们假定读者已具备数学系学生通常有的线性代数知识.因此我们要求读者熟悉向量及其分量,熟悉在已知坐标系表示线性变换的矩阵,并且要求了解等价、全等和相似的基本概念以及在这些变换下矩阵的标准形. 我们还假定读者知道关于行列式和线性方程组以及矩阵的特征值和特征向量这样一些基本知识. 关于这些内容在法捷也娃 (Фаддеева, 1950) 著的一书中的第一章已有很好的概述.

我们假定,读者熟悉在计算机上用 FORTRAN, ALGOL 60 或 PL/I 语言编制程序,并且对于数值计算及其误差已有一些了解,但对数值矩阵计算则毫不了解.

本书的目的在于使这样的读者尽快地了解求解线性方程组的一些好的程序和有关误差的基本概念.

2. 向量范数和矩阵范数

我们假定读者相当熟悉矩阵 (用来表示线性变换的) 代数. 由于在计算中直接涉及各种数的大小,所以要理解数值矩阵方法,还必须稍为懂得一点矩阵几何和矩阵分析. 在这章和下一章我们将介绍这些概念. 带有证明的更详细的讨论可参看法捷也娃的书,但在该书中没有定理 (3.1).

令 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T$ 表示在实 n 维空间 R^n 中的列向量. $\boldsymbol{x}$ 的转置 $\boldsymbol{x}^T$ 表示行向量. 我们引入向量 $\boldsymbol{x}$ 的欧几里得 (Euclid) 长度或范数的概念,并用 $\|\boldsymbol{x}\|$ 表示,定义为:

(2.1) $\|\boldsymbol{x}\| = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$.

在二维或三维空间中的范数具有普通长度的性质:

(2.2) $\|c\boldsymbol{x}\| = |c| \cdot \|\boldsymbol{x}\|$. 对所有实数 c 和所有向量 $\boldsymbol{x}$.

(2.3) $\|\boldsymbol{0}\| = 0$ 和 $\|\boldsymbol{x}\| > 0$. 如果 $\boldsymbol{x} \neq \boldsymbol{0}$.

(这里 $\boldsymbol{0}$ 表示零向量.)

(2.4) $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, 对所有向量 $\boldsymbol{x}, \boldsymbol{y}$.

在 (2.4) 中当且仅当 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 线性相关, 即它们是共线向量时等号才成立.

性质(2.2)和(2.3)可直接从 (2.1) 得出,(2.4)的证明只用到初等分析,我们留给读者. 为证(2.4),可首先证明著名的哥西-许瓦兹-布尼亚可夫斯基 (Cauchy-Schwarz-Буняковский)不等式,

(2.5) $|\boldsymbol{x}^T\boldsymbol{y}| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$. 对任何 $\boldsymbol{x}, \boldsymbol{y}$.

注意到实变量 α, β 的二次函数

$$(\alpha\boldsymbol{x} + \beta\boldsymbol{y})^T(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha^2\|\boldsymbol{x}\|^2 + 2\alpha\beta\boldsymbol{x}^T\boldsymbol{y} + \beta^2\|\boldsymbol{y}\|^2$$

非负的必要条件,即可证明不等式 (2.5).

我们现在定义 n 行和 n 列实矩阵 A 的范数为

(2.6) $$\|A\| = \max_{\boldsymbol{x}\neq 0} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}.$$

由定义直接得出

(2.7) $\|cA\| = |c| \cdot \|A\|$. 对所有实数 c 和所有 A,

(2.8) $\|\Theta\| = 0$; 若 $A \neq \Theta$ 则 $\|A\| > 0$.

(这里 Θ 表示零矩阵.)

(2.9) $\|A + B\| \leqslant \|A\| + \|B\|$. 对所有 $n \times n$ 矩阵 A 和 B,

这些性质与向量范数的性质 (2.2), (2.3), (2.4) 完全相同. 因此可把 $n \times n$ 矩阵 A 的集合看成是 n^2 维赋范向量空间.

从定义 (2.6) 还可直接得出,

(2.10) $\|A\boldsymbol{x}\| \leqslant \|A\| \cdot \|\boldsymbol{x}\|$, 对所有 $A, \boldsymbol{x}$,

并且，不等式 (2.10) 总是"精确的"或"最佳的"，即对每个矩阵 A，存在向量 $\boldsymbol{x}$，使得

$$(2.11)\qquad \|A\boldsymbol{x}\| = \|A\| \cdot \|\boldsymbol{x}\|.$$

读者从 (2.6) 可证明

$$(2.12)\qquad \|AB\| \leqslant \|A\| \cdot \|B\|,\ \text{对所有矩阵 } A,\ B.$$

有了性质 (2.10) 和 (2.12) 就使得向量和矩阵的范数在研究线性变换，特别是在分析解线性方程组的误差时非常有用.

因为方阵 A 表示一个 n 维空间 X 的每个向量到另一个 n 维空间 Y 中的向量 $\boldsymbol{y} = A\boldsymbol{x}$ 的线性变换 (映象). 所以定义 (2.6) 可等价地表示为

$$(2.13)\qquad \|A\| = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\|,$$

从而可把 $\|A\|$ 解释为单位球面 $\{\boldsymbol{x}: \|\boldsymbol{x}\| = 1\}$ 在变换 $\boldsymbol{x} \to A\boldsymbol{x}$ 下的象集 $\{A\boldsymbol{x}\}$ 中最长向量的长度.

回想正交矩阵 U 具有特性 $U^TU = UU^T = I$，其中 I 是单位矩阵. 变换 $\boldsymbol{x} \to U\boldsymbol{x}$ 表示一个 n 维空间自身的刚性旋转，但可能关于某个超平面预先作一次反射.

欧几里得范数相当于二维和三维空间中普通长度的概念，这使我们在高维时可以利用几何直观. 在正交矩阵作用下这个长度保持不变，即对所有正交矩阵 U，$\|\boldsymbol{x}\| = \|U\boldsymbol{x}\|$.

我们只讨论了一种范数，其实可能定义的向量范数有无穷多. 在数值分析中经常使用的另外两种范数是

$$(2.14)\qquad \|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

和

$$(2.15)\qquad \|\boldsymbol{x}\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|.$$

这些向量范数满足 (2.2)，(2.3)，(2.4)，并且由 (2.6) 可导出相应的新的矩阵范数. 例如在法捷也娃的书中已证明有:

$$(2.16)\qquad \|A\|_1 = \max_{1\leqslant j\leqslant n} \sum_{i=1}^{n} |a_{i,j}|,$$

$$(2.17)\qquad \|A\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |a_{i,j}|.$$

这两种范数不允许随意运用几何直观，并且他们对应的保范矩阵通常不是正交矩阵． 然而范数(2.16)和(2.17)比欧几里得范数容易计算,在许多应用中是很有用的.

本节几乎所有的叙述都可以搬到复向量和复矩阵．只须用复共轭转置 $\boldsymbol{x}^H = \bar{\boldsymbol{x}}^T$ 和 $A^H = \bar{A}^T$ 代替 $\boldsymbol{x}^T$ 和 A^T 即可.

(2.18) **练习**．证明(2.4),(2.12),(2.16)和(2.17).

(2.19) **练习**．证明对任何 $n\times n$ 矩阵 A，有

$$\max_{i,j}|a_{i,j}| \leqslant \|A\| \leqslant n\max_{i,j}|a_{i,j}|.$$

3. 在正交等价变换下矩阵的对角形

现在我们来叙述由矩阵 A 表示的从一个 n 维欧几里得空间到另一个 n 维空间的线性变换的性质的某些几何直观．假定在两个空间中已给定正交坐标系，作者认为，下述定理(3.1)给出由矩阵表示的线性变换的最简单的解释．注意正交矩阵已在第2章定义.

(3.1) **定理**　给定任一 $n\times n$ 实矩阵 A，存在两个 $n\times n$ 实正交矩阵 U，V，使得 U^TAV 是对角矩阵 D．而且,我们可以选择 U 和 V，使得 D 的对角元素是

$$\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_r > \mu_{r+1} = \cdots = \mu_n = 0,$$

其中 r 是 A 的秩．特别,若 A 是非奇异的,则

$$\mu_1 \geqslant \mu_2 \geqslant \cdots > \mu_n > 0.$$

数 μ_1，$\cdots$，μ_n 是 A 的奇异值．如在本章末将指出的,它们是

对称矩阵 AA^T 的特征值的非负平方根（必然非负），这里 A^T 是 A 的转置矩阵．实际上，求解矩阵 A 的方程组的一切重要事实都同 A 的奇异值集合的特性有关．(3.1) 的证明将在第 4 章给出．

在线性代数教科书里，通常是用一般的非奇异矩阵 P^{-1}，Q 代替 U^T，V，而用 1 代替所有的 μ_1，$\cdots$，μ_r 来叙述定理 (3.1) 的．但从计算观点看，正交矩阵 U，V 有更大的价值，因为对任何正交矩阵 U 和任何向量 $\boldsymbol{x}$，$\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$．即用正交矩阵相乘不改变所涉及数的大小．而用一般非奇异矩阵 P 相乘则可能发生剧烈的变化．这是一个重要的区别．

要理解定理 (3.1) 为何如此有用，我们来考虑由矩阵 A 表示的从一个 n 维空间 X 到另一个 n 维空间 Y 的线性变换．这样，对 X 中的任何 $\boldsymbol{x}$，$\boldsymbol{y} = A\boldsymbol{x}$ 属于 Y．在用矩阵 A 表示线性变换时，我们曾假定在 X 与 Y 中已给定了正交坐标系．对 X 中的坐标系作一正交变换，于是用 $\boldsymbol{x}$ 表示的向量具有新的表示 $\boldsymbol{x}'$，其中 $\boldsymbol{x} = V\boldsymbol{x}'$．同样，在 Y 中用不同的正交坐标变换，我们得到 y 的新坐标 $\boldsymbol{y}'$，其中 $\boldsymbol{y} = U\boldsymbol{y}'$．这里 U 和 V 都是 (3.1) 中的矩阵．

由于在 X 和 Y 中基底变换的结果，原来用 A 表示的变换现在有新的表示，我们将证明它就是 D．事实上，

$$\boldsymbol{y}' = U^T\boldsymbol{y} = U^T A\boldsymbol{x} = U^T A(V\boldsymbol{x}') = (U^T AV)\boldsymbol{x}' = D\boldsymbol{x}'.$$

即 $\boldsymbol{y}' = D\boldsymbol{x}'$．这就是要证明的．

在新的正交坐标系中，变换显得十分简单．用分量形式表示就是

$$\begin{cases} y_1' = \mu_1 x_1' \\ y_2' = \mu_2 x_2' \\ \vdots \end{cases}$$

$$(3.2)\qquad \begin{cases} y'_r = \mu_r x'_r \\ y'_{r+1} = 0 \\ \vdots \\ y'_n = 0. \end{cases}$$

上述变换只不过是一个伸缩变换，它以伸缩因子 $\mu_1 > 0$ 将 X 中的第一个坐标轴映射到空间 Y 的第一个坐标轴，同样，以伸缩因子 $\mu_2, \cdots, \mu_r$ 将对 X 中的第 2 个，第 3 个，……，第 r 个坐标轴映射到 Y 中的第 2 个，……，第 r 个坐标轴．而 X 中的第 $r+1$ 个……，第 n 个坐标轴映射到 Y 中为零向量．

令 A^T 是 A 的转置，于是

$$D^TD = (U^TAV)^T(U^TAV) = V^TA^TUU^TAV = V^TA^TAV.$$

因此，$V^T(A^TA)V = D^TD$，它是一个对角元素为 $\mu_1^2, \cdots, \mu_r^2, 0, \cdots, 0$ 的对角矩阵．因为 V 正交，$V^T = V^{-1}$，而且变换 $V^T(A^TA)V$ 不改变 A^TA 的特征值．A^TA 的特征值是 $\mu_1^2, \cdots, \mu_n^2$．所以 A 的奇异值是 A^TA 的特征值的非负平方根．最后这一性质常用以确定这些奇异值．

从 (3.2) 我们可以证明 D 将单位球面 $S = \{\boldsymbol{x}': \|\boldsymbol{x}'\| = 1\}$ 映射为向量 $\boldsymbol{y}'$ 的 r 维超椭球面 $E = DS$：

$$\frac{y_1'}{\mu_1^2} + \cdots + \frac{y_r'^2}{\mu_r^2} = 1 \quad 且 \quad y'_{r+1} = \cdots = y'_n = 0.$$

E 中距原点 θ 最远的点是 $(\mu_1, 0, \cdots, 0)$．若 $r < n$，则 E 包含原点 θ．若 $r = n$，则 E 不包含原点并且 E 中最接近于 θ 的点是 $(0, \cdots, 0, \mu_n)$．若 $r < n$，则 D，因而 A 是奇异矩阵．若 $r = n$，D 和 A 为非奇异因而有逆矩阵；因此从 (3.2) 直接看出

$$D^{-1} = \begin{bmatrix} \mu_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \mu_n^{-1} \end{bmatrix}.$$

所以 A^{-1} 的奇异值是 $\mu_1^{-1}, \cdots, \mu_n^{-1}$.

从上面的讨论和 (2.6) 我们看到

(3.3) $$\|A\| = \|D\| = \mu_1.$$

若 $r = n$, 则

(3.4) $$\|A^{-1}\| = \|D^{-1}\| = \mu_n^{-1}.$$

于是有奇异值 $\mu_1 \geqslant \cdots \geqslant \mu_n > 0$ 的非奇异方阵 A 的最重要之点可概括为: 在 X 中有直线 L_1, 当 L_1 映象为 Y 中直线 AL_1 时, 矩阵 A 将 L_1 伸长(或压缩) μ_1 倍, 与 L_1 正交的第二条直线 L_n, 矩阵 A 将 L_n 伸长(或压缩) μ_n 倍, 并且 AL_1 和 AL_n 在 Y 中正交. 矩阵 A 将平面 L_1, L_n 上的单位圆映射为具有半轴 μ_1 和 μ_n 的椭圆. 这是 X 中任何圆可能出现的最大畸变.

注意 A 的行列式 $\det(A)$ 满足条件

$$|\det(A)| = \mu_1\mu_2\cdots\mu_n.$$

证明可由定理 (3.1), 等式

$$\det(A) = \det(U^T)\cdot\det(D)\cdot\det(V)$$

和正交矩阵的行列式是 1 或 -1 得出.

(3.5) **练习.** 若 A 是对称矩阵, 证明在定理 (3.1) 中奇异值 μ_i 是在特定的次序下的数 $|\lambda_i|$, 其中 λ_i 是 A 的特征值.

4. 对角形定理的证明

为证定理 (3.1), 注意 $B = AA^T$ 是对称矩阵. 对所有向量 $\boldsymbol{x}$,

$$\boldsymbol{x}^T B\boldsymbol{x} = \boldsymbol{x}^T AA^T\boldsymbol{x} = (A^T\boldsymbol{x})^T(A^T\boldsymbol{x}) = \boldsymbol{y}^T\boldsymbol{y} = \|\boldsymbol{y}\|^2 \geqslant 0,$$

这里 $\boldsymbol{y} = A^T\boldsymbol{x}$, 可见 B 是半正定矩阵. 从而它的 n 个特征值都是非负的, 所以可令它们是 $\mu_1^2, \mu_2^2, \cdots, \mu_n^2$, 这里

$$\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n \geqslant 0.$$

假定 $\mu_r > 0$，或者 $r = n$，或者 $\mu_{r+1} = \cdots = \mu_n = 0$.

回想实对称矩阵的特征值理论，我们可找到正交矩阵 U 使得

$$U^T B U = D^2,$$

其中 D 是对角矩阵，其对角元素是

$$d_{i,i} = \mu_i \geqslant 0 \quad (i = 1, 2, \cdots, n).$$

现在定义 $n \times n$ 矩阵 F 为

$$F = U^T A, \tag{4.1}$$

则有 $FF^T = (U^T A)(A^T U) = D^2$，即

$$(FF^T) = D^2, \tag{4.2}$$

它是对角矩阵. 等式(4.2)的第 (i, i) 个元素表明，F 的第 i 行 $\boldsymbol{f}_i$ 的范数是 $\mu_i (i = 1, 2, \cdots, n)$. 即 $\|\boldsymbol{f}_i\| = \mu_i$. 且(4.2)的非对角元素等于零表明 F 的不同行是相互正交的. 因为 $\mu_1 \geqslant \mu_2 \geqslant \cdots \mu_r > 0$，所以 F 的前 r 行是相互正交的非零向量 $\boldsymbol{f}_1, \cdots, \boldsymbol{f}_r$. 若 $r < n$，则剩下的行 $\boldsymbol{f}_{r+1}, \cdots, \boldsymbol{f}_n$ 是零向量 $\mathbf{0}$.

现在，我们用以下的方法构造行向量 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 的正交组:

(4.3) 对 $i = 1, \cdots, r$，令 $\boldsymbol{v}_i = (1/\mu_i) \cdot \boldsymbol{f}_i$.

因此，$\|\boldsymbol{v}_1\| = \cdots = \|\boldsymbol{v}_r\| = 1$. 对 $i = r + 1, \cdots, n$，选择范数为 1 的向量 $\boldsymbol{v}_i$，使 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 相互正交. 在线代数中已证明这样构造正交基的过程总是可能的.

令 V^T 表示这样的矩阵，它的行向量是 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$:

$$V^T = \begin{bmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_n \end{bmatrix}.$$

那末，向量 $\{\boldsymbol{v}_i\}$ 的正交性意味着 $V^T V = I$，所以 V 是正交矩阵. 而且，从(4.3)和以上所述可得

$$f_i = \mu_i v_i \quad (i = 1, \cdots, n).$$

因此

(4.4) $$F = DV^T.$$

但从(4.1)和(4.4)可知

$$U^T A = DV^T,$$

因此

$$U^T A V = D,$$

从而定理(3.1)得证.

定理(3.1)稍加改变就可以推广到任意长方矩阵 A. 为完备起见,我们叙述这个结果:

(4.5) **定理**. *给定任何秩为 r 的 n 行 k 列实矩阵 A, 则存在 $n \times n$ 实正交矩阵 U 和 $k \times k$ 实正交矩阵 V, 使得 $U^T A V$ 是 n 行 k 列形如*

$$D = \begin{bmatrix} \mu_1 & & & 0 \\ & \mu_2 & & \\ & & \ddots & \\ 0 & & & \mu_r \end{bmatrix}$$

的矩阵. 其中 $\mu_1 \geqslant \mu_2 \cdots \geqslant \mu_r > 0$.

(4.6) **练习**. 证明(4.5). 提示: 仿效(3.1)的证明,

令

$$U^T A A^T U = \begin{bmatrix} D & \Theta \\ \Theta & \Theta \end{bmatrix} \quad 和 \quad U^T A = \begin{bmatrix} F \\ G \end{bmatrix}.$$

证明 $G = \Theta$, 等等.

(4.7) **练习** 给定任意非奇异矩阵 A, 求与它最接近的奇异矩阵的距离, 即求 $\|A - S\|$ 关于奇异矩阵 S 的极小值. (答案: μ_n). 最近的奇异矩阵是唯一的吗?

(4.8) **练习**. 给定任意奇异矩阵 A, 求与它最近的非奇异矩阵的距离.