增刊 第1期 1985年8月

自 动 化 学 报
ACTA AUTOMATICA SINICA

11卷

用奇异值分解实现捷联式矩阵最佳正归化的方法

毛剑琴
（北京航空学院）

摘 要

本文提出了一种矩阵最佳正归化*的新方法——奇异向量法，并论述了这一方法的理论依据。通过与其它方法在收敛性、精度及计算次数方面的对比，说明了奇异向量法的优越性。

一、前 言

通过数值计算实现矩阵的最佳正归化是数值代数中的一个问题，也是捷联式惯导系统算法中的一个重要问题。十几年来，一些学者对这一具有理论意义及实用价值的问题进行了研究，得到了有价值的结果。本文将以捷联式惯导系统中的方向余弦阵为具体对象来讨论矩阵最佳正归化的一般方法。

在捷联式惯导系统中，机体系与平台系之间的方向余弦阵（DCM），即捷联式矩阵的计算是确定飞行器姿态的关键。通常由方向余弦法或四元数法求得捷联式矩阵。

捷联式矩阵是两个笛卡尔坐标系之间的变换，因此它应是一个正交归一的矩阵。但由于舍入误差、量化误差、解微分方程中的算法误差等的存在，往往不可能得到精确解，即由计算得到的捷联式矩阵是不正归的。

在数值计算与精确解之间的误差中有一部分是可以通过数值方法进行修正的，称之为非正归化误差。如何通过数值方法消除非正归化误差是本文讨论的重点。数学的提法是，对一个非正归矩阵，如何用数值的方法找到一个正归矩阵，使其二者的距离为最小。这就是所谓矩阵的最佳正归化问题。

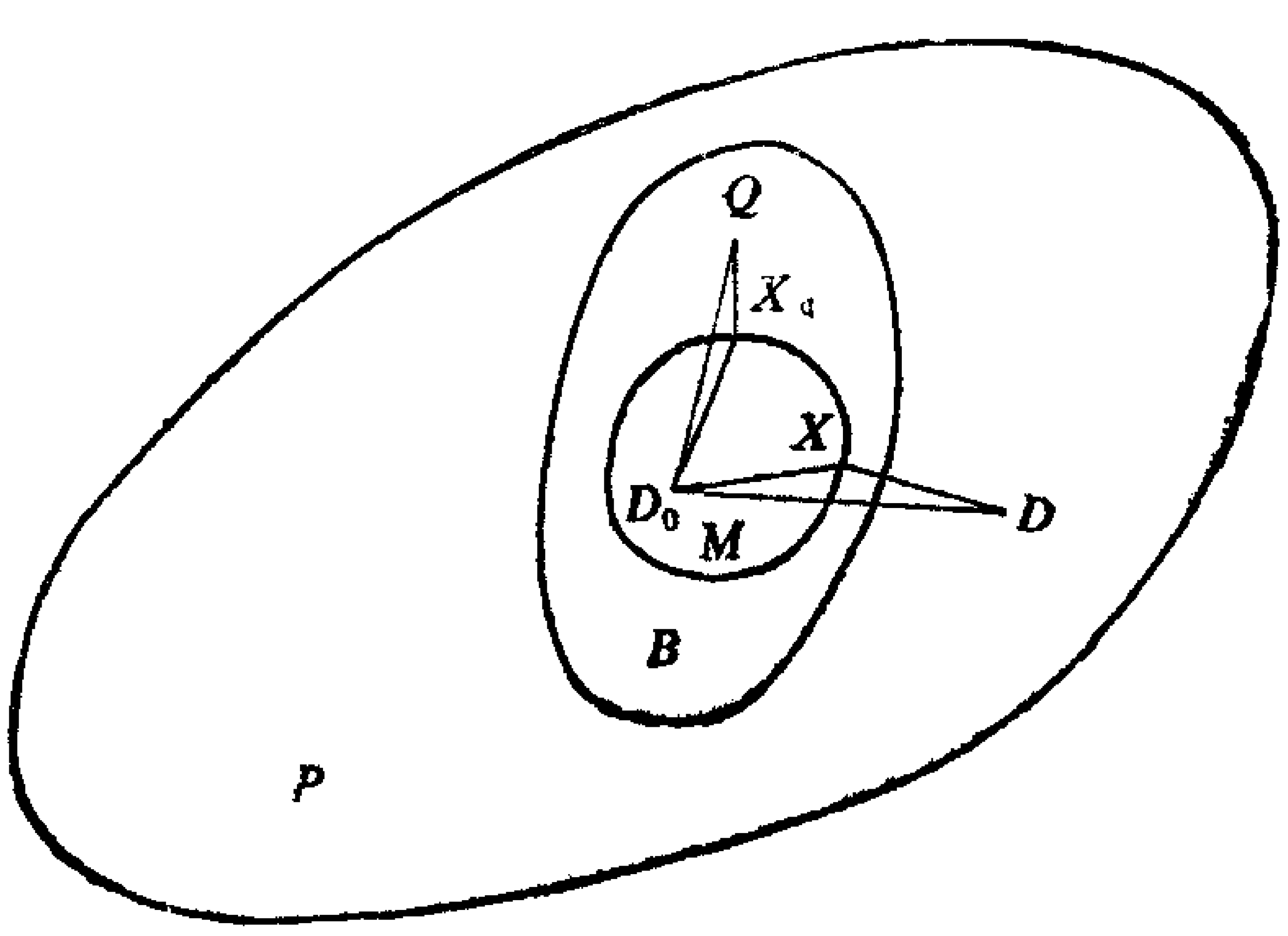

图 1 捷联式矩阵数值解、精确解及正归化解之间的关系

四元数的情形下，将产生正交但不归一的捷联式矩阵，其归一化问题是一般捷联式矩阵

本文于1983年12月1日收到。

* 所谓“正归”(或“正交归一”) 即数学中的规格化正交。“正交”是指矩阵各列(或行)向量相互垂直;“归一”是指各列(或行)向量的模为一。这样的矩阵在惯导中常称为“正归阵”。

正归化问题的特殊情形．这几个解的集合见图1．图中M为正归矩阵集合，B为正交矩阵集合，P为任意矩阵集合．D_0为精确解，Q为通过四元数所求得之数值解，D为通过方向余弦法所求得之数值解．X_q为Q的归一化矩阵，X为D的最佳正归化矩阵．数值解D与精确解D_0之间的误差为$\overline{D_0D}$，经过最佳正归化修正得到矩阵X，这时必定有$\overline{D_0X}<\overline{D_0D}$，即通过最佳正归化，消除了一部分误差．对于Q，$X_q$，$D_0$也有类似的关系．

下面要讨论的问题是如何找到精度高、数值稳定的方法来求得最佳正归化矩阵X．

二、矩阵最佳正归化的迭代方法和奇异向量法

近十几年来，I. Y. Bar-Itzhack [1-7]，C. R. Giardina [8] 等和A. Björk [9] 等在这方面的工作具有代表性．现简单介绍如下．

设非正归阵D为由方向余弦法得到的数值解，现要求正归阵X，使其二矩阵在欧氏模的意义下距离最小，即设$\bar{X}$为任意正归阵，则X将使

$$N\{D-\bar{X}\}\triangleq(\Sigma_i\Sigma_j(D_{ij}-\bar{X}_{ij})^2)^{1/2} \tag{1}$$

达到最小，或等价地使

$$N\{D-\bar{X}\}\triangleq(t_r[(D-\bar{X})^T(D-\bar{X})])^{1/2} \tag{2}$$

达到最小。

求X的近似解的迭代方法有三种：

1）Björk和Bowie迭代法 [9]

$$\begin{cases} X_0=D, \\ X_{n+1}=\dfrac{3}{2}X_n-\dfrac{1}{2}X_nX_n^TX_n. \end{cases}$$

2）双迭代法 [3,4]

$$\begin{cases} X_0=D, \\ X_{n+1}=\dfrac{1}{2}(X_n^T)^{-1}+\dfrac{1}{2}X_n. \end{cases}$$

3）梯度投影法 [1]

$$\begin{cases} X_0=D, \\ X_{n+1}=X_n-\dfrac{1}{2}X_nX_0^TX_n+\dfrac{1}{2}X_0. \end{cases}$$

文〔2〕中给出了为保证收敛性，初始矩阵$X_0=D$与最佳正归矩阵的偏差须保持在一定的范围内迭代收敛，否则将发散．因而迭代法在实际应用中是不够理想的，有必要寻找一种数值上更加稳定的方法．

现引入矩阵奇异值分解的概念．

定理1. [11] 设$A\in R^{m\times n}$，则存在正归矩阵$U\in R^{m\times m}$，正归矩阵$V\in R^{n\times n}$及$\Sigma\in R^{m\times n}$，使得

$$A = U\Sigma V^T.$$

其中，

$$\Sigma = \begin{pmatrix} S \\ 0 \end{pmatrix},$$

$$S = \mathrm{diag}\ (\sigma_1, \cdots\cdots, \sigma_n),$$

并有

$$\sigma_1 \geqslant \sigma_2 \geqslant \cdots\cdots \geqslant \sigma_n \geqslant 0.$$

$\sigma_i\,(i=1,\cdots\cdots,n)$ 为矩阵A的奇异值，U与V分别为矩阵A的左奇异向量阵与右奇异向量阵。A的这种分解称为奇异值分解。

对于捷联式矩阵$D\in R^{3\times3}$，存在正归阵 $U\in R^{3\times3}$，$V\in R^{3\times3}$及对角阵$\Sigma\in R^{3\times3}$，使$D=U\Sigma V^T$。

引理[8]．若D为3×3任意实矩阵，X为正归矩阵，则要X使欧氏模$\|D-\tilde{X}\|$取极小值，当且仅当X使$t_r(D\tilde{X}^T)$取极大。

在此基础上，可以证明下述定理，它是矩阵最佳正归化的奇异向量法的理论基础。

定理2.　设D为3×3实矩阵，其秩至少为2，则使（1）或（2）式取最小值的最佳正归化矩阵为

$$X = UV^T.$$

其中，U、V为D的奇异向量阵，即

$$D = U\Sigma V^T.$$

证明．1）若D非奇异，并有$D=U\Sigma V^T$．设X为正归阵，并令$\tilde{X}=U^{-1}XV^{-T}$，则$X=U\tilde{X}V^T$。于是利用矩阵迹的性质[10]及U，V正归的性质，有

$$\begin{aligned} t_r(DX^T) &= t_r(U\Sigma V^T\cdot V\tilde{X}^TU^T) \\ &= t_r(\Sigma\tilde{X}) \\ &= \tilde{X}_{11}\sigma_1 + \tilde{X}_{22}\sigma_2 + \tilde{X}_{33}\sigma_3. \end{aligned} \tag{3}$$

又因为$\tilde{X}$为正归阵，所以$\tilde{X}_{i1}^2+\tilde{X}_{i2}^2+\tilde{X}_{i3}^2=1\ (i=1,2,3)$，即$\tilde{X}_{ii}^2\leqslant 1\ (i=1,2,3)$。由于$\tilde{X}_{ij}\ (i\neq j)$不出现在$t_r(DX^T)$的展开式中，并且$\sigma_i\geqslant 0$，若要使$t_r(DX^T)$达到最大，当且仅当

$$\begin{cases} \tilde{X}_{ij}=1\ (i=j), \\ \tilde{X}_{ij}=0\ (i\neq j), \end{cases} \tag{4}$$

所以$\tilde{X}=I$，即$X=UV^T$。

2）若D为奇异，其秩为2，此时$\sigma_3=0$。由（3）、（4）式有$\tilde{X}_{11}=\tilde{X}_{22}=1$及$\tilde{X}_{12}=\tilde{X}_{13}=\tilde{X}_{21}=\tilde{X}_{23}=0$。又由于X为正归阵，则可取$\tilde{X}_{33}=1$及$\tilde{X}_{31}=\tilde{X}_{32}=0$。于是得到同样的结果$\tilde{X}=I$，$X=UV^T$。证完。

推论1.　在四元数情形下，Q为正交不归一阵，有奇异值分解为$Q=U_q\Sigma_qV_q^T$，其中U_q，V_q为正归阵，$\Sigma_q=|\tilde{q}|\cdot I$，所以$X_q=Q/|\tilde{q}|$。

推论2.　上述定理可推广到$D\in R^{n\times n}$的情形，但当$n\geqslant 3$时，最佳正归化阵X可能不唯一。

推论3.　矩阵D的奇异值分解一般不唯一，即除了Σ唯一确定外，U，V不唯一确定。然而，不难证明UV^T却是唯一确定的，因此X是唯一确定的。

按文〔12〕计算矩阵奇异值分解．可以证明奇异向量法的相对误差为

$$\|\tilde{U}\tilde{V}^T - UV^T\|_F \leqslant \frac{2}{\sigma_n}\|\tilde{D} - D\|_F$$

其中$\tilde{D} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$是数值计算后D的实际值；$\sigma_n$是D的最小奇异值。所以当$\sigma_n$足够大时，奇异向量法具有很好的数值稳定性，且其精度与计算机的精度至多相差一到两个量级。

综上所述，与迭代法相比，奇异向量法的优点是保证了收敛性和计算精度。值得指出的是，这些优点是用增加了算法的复杂性得来的，因此它将占用更多的内存。

三、数值计算例子

在NOVA-4机上，取双精度字长为64位，用LINPACK数学库中的标准子程序SSVDC对文〔7〕中的例子进行了计算与比较，其结果如下（迭代法的结果取自文〔7〕）：

例1.

$$D = \begin{pmatrix} 0.40735173 & -0.80419803 & 0.11052590 \\ -0.88363382 & -0.77214510 & -0.54520913 \\ -0.90991876 & 0.75857107 & -0.86116686 \end{pmatrix},$$

$$X = \begin{pmatrix} 0.61489492 & -0.59950310 & -0.51234781 \\ -0.74938374 & -0.64654371 & -0.14284691 \\ -0.24561809 & 0.47178095 & -0.84681432 \end{pmatrix}.$$

迭代次数	$N(X_n^TX_n - I)$			
	迭代法			奇异向量法
	1)	2)	3)	
0	2.267	2.267	2.267	0.6672×10^{-15}
⋮	⋮	⋮	⋮	
6	0.931	0.5×10^{-12}	0.451	
⋮	⋮		⋮	
13	0.2×10^{-8}		0.095	
⋮			⋮	
85			0.84×10^{-9}	

用奇异向量法求得的奇异向量阵为

$$\tilde{U} = \begin{pmatrix} -0.38553365 & -0.46431182 & -0.79735710 \\ 0.38792437 & -0.86564261 & 0.31650838 \\ 0.83718486 & 0.18728962 & -0.51385222 \end{pmatrix},$$

$$\tilde{V} = \begin{pmatrix} -0.73339464 & 0.31719380 & -0.60126566 \\ 0.37528643 & 0.92639184 & 0.03095585 \\ -0.56682661 & 0.20294398 & 0.79844933 \end{pmatrix}.$$

例2.

$$D=\begin{pmatrix} 0.33906376 & 0.36260365 & 0.29026758 \\ 0.34863198 & -0.81879170 & -0.46903664 \\ 0.81121079 & -0.36735531 & -0.93098548 \end{pmatrix},$$

$$X=\begin{pmatrix} 0.71178046 & 0.27777704 & 0.57200947 \\ 0.28205876 & -0.95575114 & 0.08356205 \\ 0.56991032 & 0.09684875 & -0.81597950 \end{pmatrix}.$$

迭代次数	$N(X_n^TX_n-I)$			
	迭代法			奇异向量法
	1)	2)	3)	
0	1.788	1.788	1.788	0.3289×10^{-15}
⋮	⋮	⋮	⋮	
6	0.6×10^{-4}	0.2×10^{-16}	0.218	
7	0.3×10^{-8}		0.155	
8	0.5×10^{-15}		0.109	
⋮			⋮	
57			0.85×10^{-8}	

用奇异向量法求出的奇异向量阵为

$$\tilde{U}=\begin{pmatrix} 0.13651046 & 0.72464609 & -0.67546498 \\ -0.59395765 & -0.48582690 & -0.64123828 \\ -0.79282987 & 0.48873332 & 0.36408863 \end{pmatrix},$$

$$\tilde{V}=\begin{pmatrix} -0.51401678 & 0.70077012 & -0.49467967 \\ 0.52881061 & 0.71295287 & 0.46049707 \\ 0.67538588 & -0.02488863 & -0.73704441 \end{pmatrix}.$$

例3.

$$D=\begin{pmatrix} -1.172399 & -1.367204 & -1.047914 \\ 1.311614 & -0.874199 & -1.499384 \\ 0.644879 & -0.992129 & 0.607769 \end{pmatrix},$$

$$X=\begin{pmatrix} -0.657449 & -0.639699 & -0.398177 \\ 0.663071 & -0.240168 & -0.708981 \\ 0.357905 & -0.730139 & 0.582064 \end{pmatrix}.$$

迭代次数	$N(X_n^T X_n - I)$			
	迭代法			奇异向量法
	1)	2)	3)	
0	5.477	5.477	5.477	0.304×10^{-15}
1	11.10	1.101	11.10	
2	246.8	0.131	283.7	
⋮	diverge	⋮	diverge	
9		0.14×10^{-5}		

用奇异向量法求出的奇异向量阵为

$$\tilde{U}=\begin{pmatrix} 0.61838575 & -0.77363630 & 0.13815185 \\ 0.77226951 & 0.56564761 & -0.28921029 \\ 0.14559832 & 0.28553399 & 0.94724415 \end{pmatrix},$$

$$\tilde{V}=\begin{pmatrix} 0.15762330 & 0.98588575 & 0.05642848 \\ -0.68736259 & 0.15056399 & -0.71053723 \\ -0.70900463 & 0.07321040 & 0.70139338 \end{pmatrix}.$$

例4.

$$D=\begin{pmatrix} 0.650865 & -1.062404 & -0.640755 \\ 0.409545 & -0.815340 & 0.208725 \\ 1.151954 & -0.621299 & -1.355879 \end{pmatrix},$$

$$X=\begin{pmatrix} -0.265287 & -0.860677 & -0.434576 \\ 0.581734 & -0.502323 & 0.639730 \\ 0.768900 & 0.083095 & -0.633946 \end{pmatrix}.$$

迭代次数	$N(X_n^T X_n - I)$			
	迭代法			奇异向量法
	1)	2)	3)	
0	4.561	4.561	4.561	0.146×10^{-15}
1	7.304	4.829	7.304	
2	55.83	0.985	127.6	
⋮	diverge	⋮	diverge	
9		0.12×10^{-5}		

用奇异向量法求出的奇异向量阵为

$$\tilde{U}=\begin{pmatrix} -0.57754258 & -0.36316838 & 0.73113152 \\ -0.24411255 & -0.77779303 & -0.57917792 \\ -0.77900811 & 0.51297829 & -0.36055463 \end{pmatrix},$$

$$\tilde{V} = \begin{pmatrix} -0.58777357 & 0.03830373 & -0.80811822 \\ 0.55496906 & 0.74590122 & -0.36829432 \\ 0.58866932 & -0.66495428 & -0.45967841 \end{pmatrix}.$$

在上述各表中，$N(X_n^T X_n - I)$表示在欧氏模意义下衡量矩阵正归化程度的指标。其中I为3×3单位阵。这一指标越小，则X_n的正归化程度越好。对于奇异向量法，表中的X_n表示由计算得到的最佳正归化阵UV^T。

由以上例子看出，迭代法由于其初始条件不同，其收敛性和收敛速度也很不相同。方法1）、3）在后两种情况下不收敛，方法2）虽然收敛，但收敛速度减慢。而奇异向量法却始终收敛并保持其精度。

最后将计算量的比较表示如下（不考虑发散情形）：

方　　法	浮点数乘、加的次数	精　　度
1）	609（6次迭代）～ 1313（或发散）（13次迭代）	$<10^{-6}\sim10^{-15}$
2）	300（5次迭代）～ 540（9次迭代）	$<10^{-5}\sim10^{-16}$
奇异向量法	524	$<10^{-15}$

由上表得知，平均地说，奇异向量法所需的计算次数并不比迭代法多。

四、结　　论

本文从理论上和数值计算上提出了矩阵最佳正归化的奇异向量法。通过数值例子的计算及与迭代法的比较，显示出奇异向量法在数值稳定性、收敛性及精度方面的优越性。

参 考 文 献

〔1〕 Itzhack Y. Bar-Itzhack, IEEE AES-**11** (1975), p.30-37.

〔2〕 Bar-Itzhack I.Y., IEEE AES-**13** (1977), p.230-235.

〔3〕 Bar-Itzhack I.Y. and Fegley K.A., IEEE AES-**5** (1969), p.798-804.

〔4〕 Bar-Itzhack I. Y., Meyer J. and Fuhrmann P. A., IEEE AES-**12** (1976), p.32-37.

〔5〕 Bar-Itzhack I.Y. and Mayer J., IEEE AES-**12**, (1976) p.146-151.

〔6〕 Bar-Itzhack I.Y., IEEE PLANS **76** (1976), p.122-127

〔7〕 Bar-Itzhack I. Y., International Journal of Numerical Methods in Engineering **11**, (1977), p.115-130.

〔8〕 Giardina C.R., Bronson R. and Wallen L., IEEE AES-**11**, (1975), p.443-446.

〔9〕 Bjǒrk A. and Bowie C. , SIAM J. , Numerical Analysis, **8**, No. 2 (1971) , p. 358–364.

〔10〕 B. D. O 安德森, J. B. 莫尔, 线性最优控制, 科学出版社 (1982年), p. 462.

〔11〕 Stewart G. W. , Introduction to Matrix Computations, Academic Press (1973) , p. 317–325.

〔12〕 Golub, G. H. and Reinsch C. , Numer. Math. , **14** (1970) , p. 403–420.

OPTIMAL ORTHONORMALIZATION OF STRAPDOW MATRIX BY SINGULAR VALUE DECOMPOSION

Mao Jianqin

(Beijing Institute of Aeronautics and Astronautics)

Abstract

In this paper, a new algorithm——singular vector algorithm for optimal orthonormalization of matrix is suggested. The advantages of this algorithm are investigated by comparing it with other iterative algorithms in convergence, accuracy and computational effort.