通过复Schur形求解线性方程 $AX+XB=-C$*

叶庆凯
(北京大学力学系,100871)

摘要:控制系统设计中经常要遇到求解线性方程

$$AX+XB=-C \tag{1}$$

的需求,因而在各种CADCS软件包中都包含求解这类方程的功能,本文介绍的方法用于矩阵设计工具MXTOOL中,其特点是适用范围广,数值稳定性高.

关键词:矩阵线性方程;复Schur形

1 引 言

考虑线性方程(1),其中 X 为 $m\times n$ 未知矩阵,A,B,C 分别为 $m\times m, n\times n, m\times n$ 已知矩阵. 该方程在稳定性理论、观测器构成等方面都很有意义.

本文提出的求解方程(1)的方法是为矩阵设计工具MXTOOL准备的. 该设计工具具有复数运算能力并提供求Hessenberg形以及进行QR迭代等常用函数.

2 相似变换下方程(1)的标准形

以酉变换矩阵 T_A, T_B 对矩阵 A,B 作相似变换,即令

$$A = T_A \times A_1 \times T_A^H, \quad B = T_B \times B_1 \times T_B^H. \tag{2}$$

其中上标 H 表示共轭转置,则方程(1)成为

$$A_1Y + YB_1 = - C_1. \tag{3}$$

其中

$$Y = T_A^H X T_B, \tag{4}$$

$$A_1 = T_A^H A T_A, \quad B_1 = T_B^H B T_B, \quad C_1 = T_A^H C T_B. \tag{5}$$

容易看出,式(3)与式(1)具有完全相同的形式.

众所周知,在允许复运算的情况下,对于一般的矩阵,总可以经过相似变换将其转换为三角形. 当方程(1)中的 A 具有下三角形 B 具有上三角形时,称其具有标准形式.

3 标准形式下方程(1)的解法

现讨论具有标准形式的方程(1). 令

$$X = [x_1\ x_2\ \cdots\ x_n], \quad C = [c_1\ c_2\ \cdots\ c_n].$$

其中 $x_i, c_i (i=1,2,\cdots,n)$ 均为 m 维列矢量. 这样,方程(1)可改写为

$$(A + b_{ii}I_m)x_i = - c_i - \sum_{j=1}^{i-1} b_{ji}x_j, \quad i = 1,2,\cdots,n. \tag{6}$$

其一般形式为

$$EY = F \tag{7}$$

---

* 国家自然科学基金资助项目.

本文于1990年12月17日收到. 1991年10月10日收到修改稿.

其中 E 是 $m\times m$ 下三角矩阵，Y，F 均为 m 维列矢量. 容易求出方程(7)的解是

$$y_i=(f_i-\sum_{j=1}^{i-1}e_{ij}y_j)/e_{ii},\quad i=1,2,\cdots,m. \tag{8}$$

显然，方程(6)有解的充要条件是矩阵 $A+b_{ii}I_m$ 的对角线上的元素均不是零，亦即方程(1)中矩阵 A 的任一特征值与矩阵 B 的任一特征值之和均不是零. 这和方程(1)有解的理论条件是完全一致的.

4 用酉变换将方阵转换为三角形的方法

众所周知，对于一个方阵 A，存在酉变换矩阵 Q，使得

$$\overline{A}=QAQ^H \tag{9}$$

具有 Schur 形.

若 A 是复矩阵，$\overline{A}$ 为复 Schur 形，即复上三角形. 若 A 是实矩阵，且限定 Q 为实的正交矩阵，则 $\overline{A}$ 为实 Schur 形. 在 A 的特征值全为实的时，$\overline{A}$ 为实上三角形，下面讨论 A 具有复特征值的情况.

设 $\overline{A}$ 的对角线上有一个 2×2 块 S，令

$$S=\begin{pmatrix}s_1 & s_2\\ s_3 & s_4\end{pmatrix}.$$

其中 $\frac{1}{4}(s_1-s_4)^2<-s_2s_3$. 可直接验证，若取酉矩阵

$$T=\frac{1}{R}\begin{pmatrix}\frac{1}{2}(s_1-s_4)-cj & s_3\\ -s_3 & \frac{1}{2}(s_1-s_4)+cj\end{pmatrix}. \tag{10}$$

其中

$$c=\sqrt{-s_2s_3-\frac{1}{4}(s_1-s_4)^2},\quad R=\sqrt{s_3^2-s_2s_3}. \tag{11}$$

则

$$TST^H=\begin{pmatrix}\times & \times\\ 0 & \times\end{pmatrix}.$$

现取一准对角形的酉矩阵，其结构形式与 $\overline{A}$ 的对角线的结构相同，且1×1块均取为1，2×2块按上述构成矩阵 T 的方法构成. 显然，用这样的酉矩阵对 $\overline{A}$ 进行酉变换可将其转换为复上三角形，即复 Schur 形.

用类似的方法可将矩阵 A 转换为复下三角形，这里不重复了.

5 算　法

综上所述，可得求解方程(1)的算法如下：

1) 将矩阵 A^H，B 转换为复 Schur 形 A_1^H，B_1，并求出相应的 T_A，T_B 使

$$A^H=T_AA_1^HT_A^H,\quad B^H=T_BB_1T_B^H.$$

2) 组成矩阵　$C_1=T_A^HCT_B$.

3)解方程(3)得到矩阵 Y；确定方程(1)的解　$X=T_AYT_B^H$.

4) 若 A,B 是实矩阵，取 X 的实部为原方程的解.

6 例　子

下面的算例都是在双精度下计算的. 为了阅读清楚起见，只写出了六位有效数字.

例 1 设 $A=\begin{pmatrix}-9 & -26 & -24\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}$, $B=A^{T}$, $C=I_3$.

由于 $$T_A=\begin{pmatrix}-0.127 & -0.471133 & 0.872872\\ -0.635001 & -0.637415 & -0.436436\\ -0.762001 & 0.609701 & 0.218218\end{pmatrix},$$

$$A_1=\begin{pmatrix}-4 & -17.2392 & 31.2888\\ 0 & -3 & 5.46101\\ 0 & 0 & -2\end{pmatrix},\quad C_1=I_3,$$

可解得 $$Y=\begin{pmatrix}0.125 & -0.307843 & 0.371661\\ -0.307843 & 1.93566 & -1.09371\\ 0.371661 & -1.09371 & 3.07803\end{pmatrix},$$

$$X=\begin{pmatrix}3.55714 & -0.5 & -0.771429\\ -0.5 & 0.771429 & -0.5\\ -0.771429 & -0.5 & 0.810119\end{pmatrix}.$$

例 2 设 $A=\begin{pmatrix}3 & 6 & 4\\ 8 & 3 & 4\\ 9 & 22 & 6\end{pmatrix}$, $B=A^{T}$, $C=I_3$.

由于

$$T_A=\begin{pmatrix}0.567351 & 0.148019-j0.75914 & 0.147268-j0.241293\\ 0.726998 & -0.006038+j0.59937 & -0.123999-j0.286211\\ 0.386764 & -0.205782-j0.013038 & 0.111035+j0.891946\end{pmatrix},$$

$$A_1=\begin{pmatrix}19.3864 & -3.74288-j1.12833 & 2.28297+j16.6874\\ 0 & -3.69322+j1.59906 & 6.40565-j1.66575\\ 0 & 0 & -3.69322-j1.59906\end{pmatrix},\quad C_1=I_3,$$

可解得

$$Y=\begin{pmatrix}-0.025791 & -0.006275-j0.001215 & 0.003627-j0.027625\\ -0.006275+j0.001215 & 0.142114 & 0.073121-j0.090953\\ 0.003627-j0.027625 & 0.073121+j0.090953 & 0.430291\end{pmatrix},$$

$$X=\begin{pmatrix}0.120267 & -0.0393082 & -0.156238\\ -0.0393082 & 0.0280664 & -0.0674334\\ -0.156238 & -0.0674334 & 0.39828\end{pmatrix}.$$

例 3 设 $A=\begin{pmatrix}3 & 6 & 4\\ 8 & 3 & 4\\ 9 & 22 & 6\end{pmatrix}^{T}$, $B=\begin{pmatrix}-9 & -26 & -24\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}$, $C=I_3$.

这里,B_1与 T_B 和例1中的 A_1与 T_A 相同,A_1与 T_A 和例2中的 A_1与 T_A 相同,且

$$C_1=\begin{bmatrix} -0.297197 & 0.0341846 & -0.954204 \\ 0.379695+j0.126511 & 0.42878+j0.803289 & -0.102899-j0.010625 \\ 0.405627-j0.766138 & -0.392888+j0.123876 & -0.140412+j0.24306 \end{bmatrix},$$

可解得

$$Y=\begin{bmatrix} -0.0193155 & -0.0112006 & -0.0972843 \\ -0.024449+j0.024687 & -0.040274-j0.12794 & 0.405156+j0.178125 \\ -0.068423+j0.098502 & 0.104536-j0.090244 & 0.377253-j0.752919 \end{bmatrix},$$

$$X=\begin{bmatrix} 0.0079846 & -0.509165 & -0.759149 \\ -0.0288551 & 0.297396 & 0.131264 \\ -0.0235652 & 0.0127866 & 0.157663 \end{bmatrix}.$$

7 结　论

本文提出的解方程(1)的方法是一种适应面很广的方法，可以解 X 是复的长方阵的情况，其唯 一的限制是方程(1)理论上应该有解存在.

本文提出的方法所需完成的计算主要是把矩阵化为复 Schur 形，这可采用一些经过广泛实践考验的算法[1]来完成，这些算法中用于变换的矩阵都是酉矩阵. 另外，本文提出的方法完全避免了数值稳定性较差的矩阵求逆运算. 因而，该方法具有既可靠数值稳定性又高的优点.

参 考 文 献

[1] 郭富印等. FORTRAN 算法汇编(第三分册)，北京：国防工业出版社，1982

Solving Linear Equation $AX+XB=-C$ via Complex Schur Form

YE Qingkai

(Department of Mechanics, Peking University·Beijing, 100871, PRC)

Abstract: In control systems design, we need to solve the matrix equation

$$AX+XB=-C.$$

Therefore, in most of the CADCS packages, there are the functions for solving this kind of equations. In this paper, a method, which is prepared for the matrix design tools MXTOOL, is introduced. It can be used for various cases, and is mathematical stable.

Key words: matrix linear equation; complex Schur form

本文作者简介

叶庆凯　1939年生. 1962年北京大学数学力学系毕业. 1979年～1981年在英国 UMIST 控制系统中心进修. 现任北京大学力学系教授. 研究兴趣为优化与最优控制中的计算方法，控制系统计算机辅助设计中的基础算法，鲁棒控制.