基于双步位移 QR 迭代的实 schur 分解实验报告

陈远洋

2025 年 1 月 17 日

1　代码构成与分析

本实验分为 my_schur.m 与 test.m 两个文件。完整代码详见相应代码文件。my_schur.m 文件主要实现了基于双步位移 QR 迭代的实 schur 分解。最后封装为 $[Q, T] = my\_schur(A)$。该分解由三步构成：

1 第一步：对输入矩阵 A 进行 Hessenberg 上三角化变为 H，矩阵右下角除次对角线外的元素为 0。

2 第二步：对上 Hessenberg 矩阵 H 进行双步位移 QR 迭代，使对角线上变为 1×1 或 2×2 块。

3 第三步：每个 2×2 块进行正交变换，使每个为实特征值的 2×2 块变为上三角阵，而有共轭复特征值的变为对角元相等的 2×2 块。

在上面三步中，第一步对应代码部分的 house_hessen 函数。第二步对应代码部分的 upper_iter 函数与 double_shift_qr 函数。第三步对应代码部分的 adjust2by2 函数。而 test.m 文件为检测部分，包含一些简单的测试样例，以及一个检查实 schur 分解是否成功的 is_schur 函数。下面是对每个函数的解析：

1. house_hessen 函数。

   输入：待求的初始矩阵 A。

   输出：Hessenberg 上三角化矩阵 H 以及该过程用到的正交变换矩阵 Q。

   利用 Householder 变换将输入矩阵 A 变换为 Hessenberg 上三角化矩阵 H。同时，记录下每一步的正交变换，计算出正交变换矩阵 Q。一共 $3n$ 次 Householder 变换，且每一次变换时间复杂度为 $O(n^2)$，故该部分时间复杂度为 $O(n^3)$。

2. upper_iter 函数

   输入：Hessenberg 上三角化矩阵 H，以及需要累积正交变换的正交矩阵 Q。

   输出：累积变换后的 Hessenberg 上三角化矩阵 H 以及累积正交变换后的 Q 子阵。

   每次迭代，首先将次对角线足够小的元素置为 0，然后分理出最大的不可约子块 $H_{l:m,l:m}$，将其送入 double_Shift_QR 函数迭代一次。并将需要更新的 Q 子阵与 H 子阵

$$(Q_{:,l:m}, H_{1:l-1:,l:m}, H_{l:m,m+1:end})$$

送入迭代过程进行累积，从而避免较大矩阵之间的乘法。以确保每次迭代时间复杂度为 $O(n^2)$。与此同时确保 $m \geq 3$，防止迭代过程出现错误。在 $m < 3$ 时，直接返回。并对已经收敛的下方保持不动。最终让 H 对角线变为 1×1 或 2×2 块。结合实际情况，迭代次数大概为 $O(n)$(可见图 3)。

3. double_shift_qr 函数

   输入：H 的不可约 Hessenberg 上三角化子阵，以及需要累积变换 H 子阵，以及需要累积正交变换的 Q 子阵。

   输出：累积变换后的 Hessenberg 上三角化矩阵 H 以及累积正交变换后的 Q 子阵。

   在 upper_iter 函数调用中可知，进行操作的 Hessenberg 子阵的一定不可约且规模大于 3×3，可以放心进行双步位移 QR 迭代。每步 householder 变换为矩阵与小规模向量更新，故时间复杂度为 $O(n)$。则总时间复杂度为 $O(n^2)$。最终期望使对角线收敛出 1×1 或 2×2 块。

4. adjust2by2 函数

   输入：已经拟上三角化的矩阵 H，以及需要累积正交变换的正交阵 Q。

   输出：实 schur 分解的矩阵 T，以及其对应的正交阵 Q。

   对输入矩阵 H，检测对角块大小。若为 2×2 块，则需要进行处理。不妨令该子块为 $A = \begin{pmatrix} x & w \\ z & y \end{pmatrix}$ 令 $\delta = |A| = xy - zw, t = tr(A) = x + y$ ，对特征值 λ 有 $\lambda^2 - t\lambda + \delta = 0$，该二次方程判别式为 $\Delta = t^2 - 4\delta$，若 $\Delta \geq 0$，则特征值为实数，否则有共轭复特征值。对实特征值情况，解出其中一个特征值 λ_1，并由 $(A - \lambda_1 I)q = 0$ 解得 q，并将其归一化 $q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$，归一化后使得尽量有 $|q_1| > |q_2|$。将其扩充为 $\begin{pmatrix} q_1 & -q_2 \\ q_2 & q_1 \end{pmatrix}$。对共轭复特征值，由 $A \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \alpha & \beta_2 \\ \beta_1 & \alpha \end{pmatrix}$ 解得对应的正交变化。其中 $\alpha = \dfrac{t}{2}, \beta_1\beta_2 = \alpha^2 - \delta$。同样的，尽量使 $|\theta| < \pi/4$。使每个为实特征值的 2×2 块变为上三角阵，而有共轭复特征值的变为对角元相等的 2×2 块。容易知道，该部分时间复杂度为 $O(n^2)$。

5. is_schur 函数

   输入：实 schur 分解的矩阵 T。

   输出：分解结果 T 结果是否为拟上三角阵，并且检查对角线上为 2×2 的块对角元是否相等。如果不是输出相关信息

   is_schur 函数位于 test.m 文件中。首先设置一个认为可以忽略的阈值 ε，检测 $\|T_{i,1:i-2}\|_1$ 是否足够小。其次检测 $T_{i,i-1}$ 是否足够小, 超过阈值 ε 说明可能为 2×2 块，需进一步检测是否有子块对角元相等。可以选择是否输出认为应该是 0 元素的绝对值之和（即左下方非次对角线元素以及次对角线上非 2×2 块元素绝对值之和）。

2 数值表现

2.1 小规模表现

在 test.m 文件中，我们可以对不同矩阵 A 进行实 schur 分解。在维数较小时，我们可以用 matlab 的 schur 函数进行输出比较验证。针对 3×3 有共轭复特征值的矩阵 $\begin{pmatrix} 1 & 0 & 0 \\ 2 & \cos(\frac{\pi}{3}) & -\sin(\frac{\pi}{3}) \\ 3 & \sin(\frac{\pi}{3}) & \cos(\frac{\pi}{3}) \end{pmatrix}$。my_schur 函数的输出为：

$$\begin{pmatrix} 1.0000 & 1.8028 & -0.8345 \\ 0 & 0.5000 & 0.2315 \\ 0 & -3.2404 & 0.5000 \end{pmatrix}$$

而系统函数输出为：

$$\begin{pmatrix} 1.0000 & 0.8345 & 1.8028 \\ 0 & 0.5000 & 3.2404 \\ 0 & -0.2315 & 0.5000 \end{pmatrix}$$

可以看到，两者只是在对角线上元素的顺序不同。以及一些共轭块的位置不同。对 test.m 文件中其他矩阵进行实 schur 分解，也可以看到其输出与 matlab 的 schur 函数输出只是在对角块上元素的顺序不同。说明 my_schur 函数的输出是正确的。可以用 is_schur 函数检测分解结果是否正确。如果 schur 分解结果正确，则会有"congratulations!" 输出。如果分解结果不正确，则会有相关信息输出。几个小规模测试用例分别代表：有共轭复特征值、小规模 (2×2)、奇异、重特征值。实验中，我们都得到了"congratulations!" 输出。

2.2 Q 的正交性与分解结果的正确性

我们可以检验 Q 的正交性以及 QTQ^T 与 A 的差异。下面针对随机生成的 100×100 矩阵进行实 schur 分解。并将 $QTQ^T - A$ 与 $QQ^T - I$ 用 matlab 的 imagesc 函数进行可视化。可以看到，$QTQ^T - A$ 与 $QQ^T - I$ 的每个元素均在 10^{-15} 级别，说明 Q 的正交性较好且用 QTQ^T 刻画 A 非常精准。

2.3 迭代次数与收敛历史

更改函数接口，在 upper_iter 函数中，将迭代次数作为输出之一。对 5 到 200 阶矩阵的迭代次数做观察，我们可以发现，平均 1~2 次迭代即可收敛出一个 2×2 块或者 1×1 块。对于 $n \times n$ 矩阵，平均迭代次数为 n~2n 次左右，收敛速度较快。将迭代次数与矩阵大小关系画图，我们可以看到，随着矩阵规模的增大，迭代次数的增长保持与之相配的线性关系。同时，单独对 100×100 矩阵进行实 schur 分解，迭代次数为 100~200 次左右，画出每步迭代后已经收敛的子块阶数我们也可以看到，其与迭代次数几乎为 1:2 的关系。

2.4 T 的拟上三角性

对随机生成的 50 到 200 阶矩阵, 平均情况下左下方次对角线上非 2×2 块元素以及左下方非次对角线元素绝对值之和在 10^{-22} 级别，说明 my_schur 函数输出的 T 拟上三角情况是非常好的。

2.5 运行时效分析

对于 200 到 500 阶矩阵，采用 tic、toc 函数计算运行时间。并用 log 函数对运行时间进行对数变换，画出运行时间与矩阵规模的关系。可以看到，my_schur 函数的运行时间在 $O(n^3)$ 级别。且对于在 200 阶以下的规模矩阵，平均运行时间在 1s 以内。

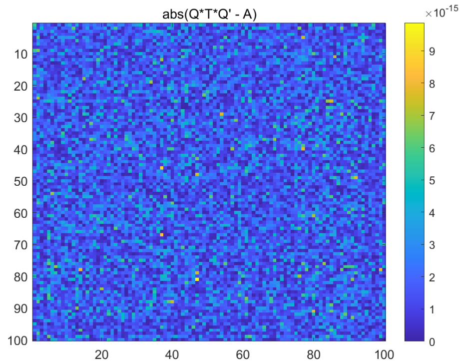

图 1: $QTQ^T - A$ 取绝对值

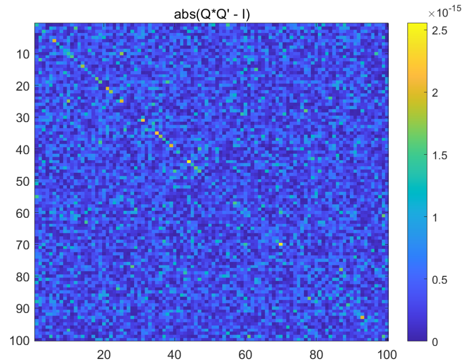

图 2: $QQ^T - I$ 取绝对值

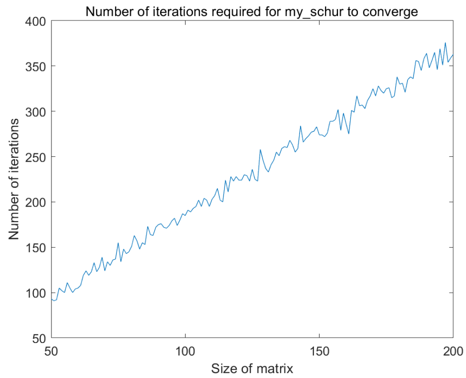

图 3: QR 双步位移迭代次数与矩阵大小关系

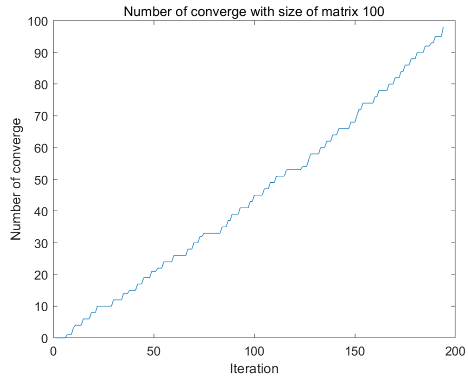

图 4: 100×100 矩阵的迭代次数与已收敛阶数关系

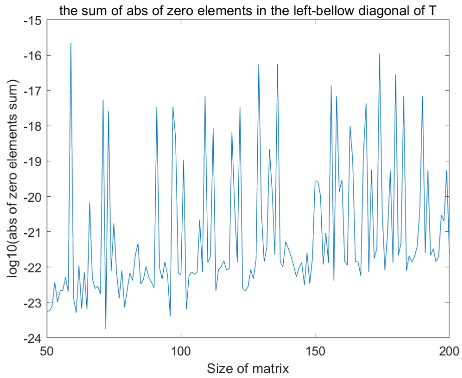

图 5: 左下角应该为的 0 元素绝对值之和与矩阵规模关系

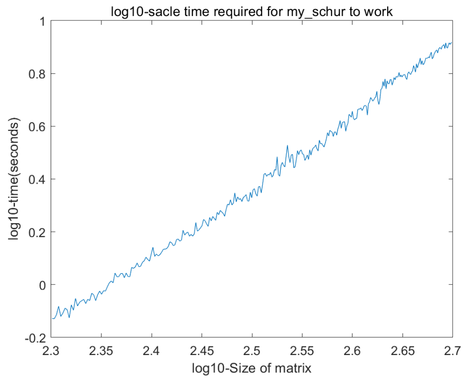

图 6: 算法运行时间与矩阵规模关系

3 总结

综上，本实验实现了基于双步位移 QR 迭代的实 $schur$ 分解。并提供了多个测试样例，用以分解结果是否正确。并从左下角零元素 (is_schur 函数)、Q 的正交性、QTQ^T 与 A 的差异等方面进行了验证。并且由于平均上只需 $O(n)$ 次双步位移 QR 迭代，故该实 $schur$ 分解过程的渐进时间复杂度为 $O(n^3)$。