Blurring of Shifts in the Multi-Shift *QR* Algorithm: Numerical Experiments using UBASIC

Roden Jason A. David
Mathematics Department
Ateneo de Manila University

ABSTRACT – The multi-shift *QR* algorithm for approximating the eigenvalues of a full matrix is known to have convergence problems if the number of shifts used in one iteration is large. The mechanism by which the values of the shifts are being transmitted from one bulge matrix to another has been discovered. In the presence of round-off errors, however, the values of the shifts are blurred in certain bulge matrices causing the *QR* algorithm to miss the true eigenvalues of the matrix. In this paper, we give the maximum number of shifts that can be used in one iteration to keep the values of the shifts from blurring. We use the UBASIC language, and specify the minimum level of precision that maintains well-focused shifts.

KEY WORDS – eigenvalues, *QR* algorithm, Schur upper triangular form, bulge-chase technique, Hessenberg form, blurring of shifts

1. The Problem and its Background

Many mathematical softwares approximate the eigenvalues of a square matrix using the *QR* algorithm. The *QR* algorithm is an iterative algorithm that approximates the Schur upper triangular form of a matrix, and the eigenvalues of the matrix emerge along the main diagonal of the Schur form. Provided that exact arithmetic is used and a convergent shift strategy is found , the QR algorithm converges quadratically for square matrices in general, and cubically for normal matrices in particular [5].

Most implementation of the *QR* algorithm use the bulge-chase technique introduced by Francis [3] in 1961. With this technique, the matrix is first reduced to upper Hessenberg form. Two iterations of the shifted *QR* algorithm are effectively performed in one *QR step*. This is done implicitly using Householder reflections that create a bulge in the matrix. The matrix is reduced back to upper Hessenberg form by chasing the bulge down the matrix. This implicit implementation of two *QR* steps in one iteration came to be known as the *double-step QR algorithm*.

In 1989, Bai and Demmel [1] generalized the double-step *QR* algorithm into the *multi-shift QR* algorithm. An arbitrary number of iterations is performed simultaneously in one multi-shift iteration. Again the same bulge chasing technique, generalized to accommodate any arbitrary number of shifts, performs this implicitly.

While multi-shift *QR* algorithm looked promising, in 1991 Dubrulle [2] presented a number of numerical experiments that show that the multi-shift *QR* algorithm does not perform very well if the number of shifts used in one multi-shift step is taken too large.

In an attempt to shed light on this convergence problem, Watkins [4] was able to identify the mechanism by which the values of the shifts are transmitted from one bulge matrix to another. His main result was that the finite eigenvalues of the matrix pencil B-μN are precisely the values of the shifts that were used in the multi-shift step. Here B is the bulge matrix and N is the Jordan nilpotent matrix of the same order as B.

Further, Watkins has detected a phenomenon called *blurring of shifts*, where the values of the shifts are incorrectly transmitted form one bulge matrix to another. Round-off errors is the apparent cause. It turns out that the characteristic polynomial of the matrix pencil B-μN, the finite zeroes of which are the shifts, is highly sensitive to perturbations of its coefficients. Since the values of the shifts are incorrectly transmitted during the bulge chase, the *QR* algorithm misses the true eigenvalues of the matrix.

2. Objectives and Methodology

In the light of the observations of Watkins [4], this study investigates the phenomenon of the blurring of shifts using extended precision arithmetic. We determine the maximum empirical number of shifts that can be used in one multi-shift step and its corresponding level of precision that keeps the shifts from blurring. Three types of matrices were studied with orders between 8 and 15 (very small). All computations were done in UBASIC allocating a maximum of 250 words per variable. All internal computations were done in 542 words per variable.

For each matrix of a given order, a set of shifts was used in the multi-shift QR algorithm and the bulge matrices were extracted from the iteration matrix. For each bulge matrix B_i, the finite zeroes of the characteristic polynomial of B_i-μN were computed and compared with the original values of the shifts. We note the maximum number *m* of shifts that can be used without exhibiting the blurring of shifts phenomenon and the minimum number of words used for the fractional part that can make this possible. This is determined as follows: beginning with a reasonably high number of worfor the fractional part, we continue to increase the number of shifts until the algorithm exhibits blurring. Having determined *m*, the maximum number of shifts, we begin to decrease the words allocated for the fractional part until the algorithm begins to fail.

The three types of matrices studied were: random matrices with integer entries taken from the set {0,1,2,…,99}, the Hilbert matrix $H=(h_{ij})$ where $h_{ij}=1/(i+j-1)$, and the tridiagonal matrix tridiag(-1,2,-1) which arises in finite-difference approximation of the second derivative.

3. Experimental Results

The tables below summarize our findings. The column labeled "Max *m*" gives the maximum number of shifts in one multi-shift iteration that were attained without exhibiting shift blurring. The column labeled "Min words" gives the corresponding minimum words per variable that can be used to attain the well-focused shifts.

Random matrix (0-99):

Size	Max *m*	Min Words
8	6	10
9	7	15
10	6	15
11	6	10
12	9	20
13	5	8
14	5	8
15	7	10

Hilbert matrix:

Size	Max *m*	Min Words
8	6	14
9	7	20
10	4	18
11	5	24
12	4	25
13	4	28
14	4	30
15	4	34

Tridiagonal:

Size	Max *m*	Min Words
8	6	6
9	7	12
10	7	10
11	6	10
12	6	12
13	5	6
14	5	6
15	5	6

We also made two observations that were not reported in the table. Firstly, some iterations start to exhibit shift blurring during the middle part of the multi-step bulge chase. For instance, with a random matrix of order 14, m=6, and 10 words for the fractional part, bulges B_0 and B_1 have well-focused shifts, while B_2 exhibits blurring. Secondly, some shift blurring become focused as the iteration progresses, but the error accumulates that the computed eigenvalues are too far from the actual values. Watkins [4] has also presented this second observation as an open question.

4. Conclusions and Recommendations

We summarize below the trends observed:

1. For matrices of order n > 12, the maximum m χ n/2. This is generally observed for matrices of higher order.
2. Hilbert matrices require greater precision for the fractional part.
3. Tridiagonal matrices require fewer words for the fractional part.
4. Some iterations exhibit blurring during the middle of the bulge chase, and some blurring become focused as the iteration progresses.

We recommend as possible future work a search for a closed formula for the maximum number of shifts in one multi-shift iteration. In the light of conclusions 1 and 2, this closed formula might be a function of the order and the condition number of the matrix. A closed formula for the minimum number of words that can be used to prevent the shifts from blurring may also be found.

References

[1] Z. Bai and J. Demmel, "On a Block Implementation of the Hessenberg Multi-Shift QR Iteration ," *Internat. J. High Speed Computing*, 1 (1989), 97-112.

[2] A. A. Dubrulle, "The Multi-Shift QR Algorithm –Is It Worth the Trouble?", *Tech. Report G320-3558x*, IBM Corp., Palo Alto, 1991.

[3] J. G. F. Francis, "The QR Transformation, parts I and II," *Computer J.* 4 (1961), 265-272, 332-345.

[4] D. S. Watkins, "Transmission of Shifts and Shift Blurring in the QR Algorithm," *Linear Algebra Appl.*, 241-243 (1996), 877-896.

[5] D. S. Watkins, and L.Elsner, "Convergence of Algorithms of Decomposition Type for the Eigenvalue Problem," *Linear Algebra Appl.,*143 (1991),19-47.