Part A Numerical Analysis, Hilary 2018. Problem Sheet 5

Note: some questions are marked optional; its up to you and your tutor if you submit these.

1. Give estimates based on Gershgorin's theorem for the eigenvalues of

$$A = \begin{bmatrix} 9 & 1 & 0 \\ 1 & 4 & \varepsilon \\ 0 & \varepsilon & 1 \end{bmatrix}, \quad |\varepsilon| < 1.$$

Find a way to establish the tighter bound $|\lambda_3 - 1| \le \varepsilon^2$ on the smallest eigenvalue of A. (Hint: consider diagonal similarity transformations.)

2. *(Optional.)* An $n \times n$ matrix has n Gershgorin discs, but when the disks overlap, there need not be an eigenvalue in each individual disc. Devise an example with $n = 2$ to illustrate this.

3. By using Gershgorin's Theorem, calculate a lower bound on the ratio $|\lambda_1|/|\lambda_2|$ where

$$|\lambda_1| \ge |\lambda_2| \ge |\lambda_3| \ge |\lambda_4|$$

are the absolute values of the eigenvalues of the matrix

$$\begin{bmatrix} -2 & 1 & 0 & 1 \\ 2 & -1 & -1 & 0 \\ 1 & 0 & 9 & -1 \\ 2 & 0 & -1 & 1 \end{bmatrix}.$$

Assume that the starting vector, x, has equal in magnitude components in all of the eigenvector directions (i.e., if $x = \sum \alpha_i v_i$ where v_i are the normalized eigenvectors, then all the α_i's are equal). Use your lower bound on the ratio to estimate how many iterations of the Power Method are needed such that the dominant term (i.e., the one in the direction of v_1) is at least 10^4 times bigger than the sum of the terms in the other eigenvector directions.

[M] Try the Power Method on the above matrix using a random stating vector. For example, using something like `y=A*x; x=y/sqrt(y'*y)` perhaps combined with a `for` loop. Stop when two successive vectors, `x`, are the same to 4 decimal places, and calculate the dominant eigenvalue as `y(j)/x(j)`. You may wish to choose `j` by selecting the biggest entry in `x`—why do you think that would matter? You can of course check if you do indeed have the dominant eigenvalue by using `eig(A)` (which uses the QR Algorithm to compute all of the eigenvalues of the matrix A.)

4. Show that all the eigenvalues of a real symmetric matrix A are real.

5. Show that if A is nonsingular then applying the Power Method with the matrix A^{-1} rather than A will give convergence to the reciprocal of the smallest eigenvalue of A in magnitude. By reference to the convergence of the Power Method, show that if the eigenvalues of A are $1, 2, 3, \ldots, n$ then convergence to the reciprocal of the smallest eigenvalue using the above approach will be much faster than convergence to the biggest eigenvalue using the Power Method with A.

Describe an efficient way to perform the Power Method with A^{-1} without explicitly computing A^{-1}.

6. Extension of the question above: if μ is known to be an approximation to an eigenvalue λ of a matrix A, why do you think it would be a good idea to apply the Power Method to the matrix $(A - \mu I)^{-1}$? Suppose that by doing so one gets convergence of `y(j)/x(j)` to γ. Show that the eigenvalue of A which one has approximately found is $\lambda = \mu + 1/\gamma$.

[**M**] Using the estimate $\mu = 9$, in the example in Question 3 above, use this procedure to calculate the biggest eigenvalue of A. Note to do this efficiently you would need to use your result from Question 5, but to just see the convergence you might consider

```
y=(A-mu*eye(4))\x; x=y/sqrt(y'*y)
```

(Recall backslash does the system solve via LU factorisation with partial pivoting.)

7. *(Optional.)* Assume that $A \in \Re^{n\times n}$ is a general matrix. If

$$J(1,n)J(1,n-1)\dots J(1,2)A = \begin{bmatrix} \alpha & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & \vdots & & \vdots \\ 0 & \times & \cdots & \times \end{bmatrix}$$

then show that $J(1,n)J(1,n-1)\dots J(1,2)AJ(1,2)^T$ is generally a full matrix (i.e., has no zeros in general) so that the result of the explicit similarity transformation

$$J(1,n)J(1,n-1)\dots J(1,2)AJ(1,2)^T J(1,3)^T \dots J(1,n)^T$$

is certainly a full matrix in general.

However, if (here we abuse notation as all of the Givens matrices are defined to make the appropriate zero based on the entries of column 1 and not column 2)

$$J(2,n)J(2,n-1)\dots J(2,3)A = \begin{bmatrix} \alpha & \times & \cdots & \times \\ \beta & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ \vdots & \vdots & & \vdots \\ 0 & \times & \cdots & \times \end{bmatrix}$$

then show that the zeros created in the first column are not destroyed by further postmultiplication by $J(2,3)^T J(2,4)^T \dots J(2,n)^T$ which makes this a similarity transformation.

Further, if A is symmetric, explain why

$$J(2,n)\dots J(2,3)AJ(2,3)^T \dots J(2,n)^T = \begin{bmatrix} \alpha & \beta & 0 & \cdots & 0 \\ \beta & \times & \times & \cdots & \times \\ 0 & \times & \times & \cdots & \times \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & \times & \times & \cdots & \times \end{bmatrix}.$$

8. *(Optional.)* Verify that in QR factorisation of a tridiagonal matrix (as in the symmetric QR Algorithm), the upper triangular matrix R has two non-zero super diagonals.

9. *(Optional.)* Verify that the sequence of matrices produced by the QR Algorithm with shifts are all similar even if a different shift if used at each iteration.

10. [M] The MATLAB command `hess` performs the explicit similarity transformation using Householder matrices described in lectures. Thus if we make a random symmetric matrix `A=randn(6,6); A=A+A'`, then `B=hess(A)` will be a tridiagonal matrix which is similar to A. Verify that `A` and `B` are indeed similar using the `eig` command.

You are now in a position to check the results of Question 8. Produce the QR factorisation of the tridiagonal matrix B. Further, check the lemma given in lectures (that tridiagonal form is preserved in the QR algorithm) by looking at where the non-zeros are in RQ.

11. *(Optional.)* [**M**] Implement the Symmetric QR Algorithm with shifting based on the last diagonal entry of a matrix. Use an dot-m file called, for example, `qr_alg_shifts` to write a function. Your code should work like:

```
n = 6;
A = randn(n,n);   A = A+A'
tol = 1e-7;
ews = qr_alg_shifts(A, tol)
ews2 = eig(A)
sort(ews) - sort(ews2)
```

You should first reduce to tridiagonal form as in Question 10. The tolerance `tol` should be used to test when the off-diagonals $b_{n-1,n}$ are sufficiently small. To *deflate* and continue on the smaller $n-1 \times n-1$ leading submatrix you can use something like:
`n = n-1; B = B(1:n,1:n);`

12. Specimen Exam question

Let A be an $n \times n$ tridiagonal matrix with the numbers $5, 10, \ldots, 5n$ on the main diagonal and the numbers ± 1 in each position $a_{j+1,j}$ and $a_{j,j+1}$. The distribution of the $\pm$ signs is arbitrary, and in particular, not necessarily symmetric.

(a) What does Gerschgorin's theorem imply about the eigenvalues of A? Give the sharpest estimates you can, and in particular, explain why although A may be non-symmetric, its eigenvalues must all be real.

(6 marks)

(b) What does it mean for a matrix A to be diagonalisable? Show that this particular matrix is diagonalisable. Show also that it is nonsingular. (5 marks)

(c) Consider the sequence of vectors $x^{(0)}, x^{(1)}, x^{(2)}, \ldots$ defined by

$$Ax^{(k+1)} = x^{(k)}, \qquad k \geq 0$$

where $x^{(0)}$ is a fixed starting vector. Prove that the vectors $x^{(k)}$ converge to 0, with the jth component of $x^{(k)}$ satisfying

$$|x_j^{(k)}| \leq 4^{-k} C$$

for some constant C. (9 marks)

(d) Suppose the recurrence is modified to

$$(A - 5I)x^{(k+1)} = x^{(k)}, \qquad k \geq 0,$$

again with a fixed starting vector $x^{(0)}$. Now what can be said about the behaviour as $k \to \infty$? Explain what algorithmic purpose might be served by carrying out this iteration. What property of $x^{(0)}$ will ensure that the recurrence achieves this purpose?

(5 marks)