

Homework 1

**Exercise 3.1** Note that by problem 4 we know that any matrix is orthonormally similar to an upper triangular matrix with the eigenvalues on the diagonal. So take  $A$ :

$$A = UTU^* \quad (1)$$

Now note:

$$\det(A) = \det(U) \det(T) \det(U^*) \quad (2)$$

$$= \det(U) \det(U^*) \det(T) \quad (3)$$

$$= \det(UU^*) \det(T) = \det(I) \det(T) = \det(T) \quad (4)$$

now remember that the determinant of an upper triangular matrix is just the product of the diagonal entries. Since the diagonal entries in this case are the eigenvalues then:

$$\det(A) = \det(T) = \prod_{i=1}^n \lambda_i \quad (5)$$

Similarly take the trace

$$\text{tr}(A) = \text{tr}(UTU^*) = \text{tr}((UT)U^*) = \text{tr}(U^*(UT)) = \text{tr}(U^*UT) = \text{tr}(T) \quad (6)$$

so once again the trace of  $T$  is just the sum of the diagonal entries which just so happens to be the eigenvalues:

$$\text{tr}(A) = \text{tr}(T) = \sum_{k=1}^n (\lambda_k) \quad (7)$$

**Exercise 3.2** to do this take an arbitrary eigenpair  $(\lambda, v)$  then note that:

$$Av = \lambda v \quad (8)$$

in particular for the largest entry  $v_i$  of  $v$ :

$$\sum_{j=1}^n A_{ij}v_j = \lambda v_i \quad (9)$$

$$\sum_{j \neq i} A_{ij}v_j = \lambda v_i - A_{ii}v_i \quad (10)$$

$$\sum_{j \neq i} A_{ij}v_j = (\lambda - A_{ii})v_i \quad (11)$$

$$(12)$$

now note that we chose  $i$  such that  $v_i > v_j$  for any other  $j$  this means that  $\frac{v_j}{v_i} < 1$  thus:

$$(\lambda - A_{ii})v_i = \sum_{j \neq i} A_{ij}v_j \quad (13)$$

$$(\lambda - A_{ii}) = \sum_{j \neq i} A_{ij} \frac{v_j}{v_i} \quad (14)$$

$$|\lambda - A_{ii}| \leq \sum_{j \neq i} |A_{ij}| \frac{v_j}{v_i} \quad (15)$$

$$| \leq \sum_{j \neq i} |A_{ij}| \quad (16)$$

$$(17)$$

So the distance from this eigenvalue to the diagonal element is less than the given sum. thus  $\lambda \in D_i$ . Since lambda was arbitrary each lambda coulda be in a different  $D_i$  but they are all in at least one by the preceding proof. Thus they are all in the union of  $D_i$ .

**Exercise 3.3** Easy proof: Because we proved it in four we know every matrix is unitarily similar to an upper triangular matrix with the eigenvalues on the diagonal. From here note that if  $A$  is skew hermitian:

$$A = -A^* \quad (18)$$

$$UTU^* = -(UTU^*)^* \quad (19)$$

$$UTU^* = -UT^*U^* \quad (20)$$

$$T = -T^* \quad (21)$$

$$(22)$$

From this we know that first of all all of the diagonal elements equal their negative conjugate (meaning that they all must be imaginary). Secondly we know that  $T$  is upper triangular but  $-T^*$  is then lower triangular as a result both of them must be diagonal since they are equal. Thus  $T$  is diagonal and  $A = UTU^*$

Longer proof since they probably want us to use decompressing the matrix:

First of all note that if  $\lambda, v$  is an eigenpair of  $A$  then:(choosing  $v$  to be unit length)

$$\lambda = \lambda \langle v, v \rangle \quad (23)$$

$$= \langle \lambda v, v \rangle \quad (24)$$

$$= \langle Av, v \rangle \quad (25)$$

$$= \langle v, A^*v \rangle \quad (26)$$

$$= \langle v, -Av \rangle = \langle v, -\lambda v \rangle = \overline{-\lambda} \langle v, v \rangle \quad (27)$$

$$= \overline{-\lambda} \quad (28)$$

So if any eigenvalue is equal to its conjugate negated, this means that it must be imaginary.

Furthermore take any matrix  $A$ . We know that it has at least one eigenvalue since it has a characteristic polynomial. Take that eigenvalue  $\lambda_1, v_1$  and corresponding unit eigenvector. construct  $Q_1 = [v_1, q_2, \dots, q_n]$  then note that:

$$Q_1^* A Q_1 = Q_1 \begin{bmatrix} A v_1 & A q_2 & \dots & A q_n \end{bmatrix} \quad (29)$$

$$= Q_1 \begin{bmatrix} \lambda_1 v_1 & A q_2 & \dots & A q_n \end{bmatrix} \quad (30)$$

$$= \begin{bmatrix} \langle \lambda_1 v_1, v_1 \rangle & \langle A q_2, v_1 \rangle & \dots & \langle A q_n, v_1 \rangle \\ \langle \lambda_1 v_1, q_2 \rangle & \langle A q_2, q_2 \rangle & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle \lambda_1 v_1, q_n \rangle & \dots & \dots & \langle A q_n, q_n \rangle \end{bmatrix} \quad (31)$$

$$\begin{bmatrix} \langle \lambda_1 v_1, v_1 \rangle & \langle A q_2, v_1 \rangle & \dots & \langle A q_n, v_1 \rangle \\ 0 & \langle A q_2, q_2 \rangle & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \langle A q_n, q_n \rangle \end{bmatrix} \quad (32)$$

$$(33)$$

Note from there that this new matrix  $Q_1^* A Q_1$  is also skew hermitian  $-A^* = -(Q_1^* A Q_1)^* = -Q_1^* A^* Q_1 = Q_1^* (-(-A)) Q_1 = Q_1^* A Q_1$ . Thus this lower block  $A_1$  is also skew hermitian. and the upper right part must be zero:

$$\begin{bmatrix} \langle \lambda_1 v_1, v_1 \rangle & 0 & \dots & 0 \\ 0 & \langle A q_2, q_2 \rangle & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \langle A q_n, q_n \rangle \end{bmatrix} \quad (34)$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \quad (35)$$

Furthermore we can repeat this process again. Note that  $A_1$  thus has at least one eigenpair  $v_2, \lambda_2$  construct an orthonormal basis  $U_2 = [v_2, z_2, \dots, z_{n-1}] \in \mathbb{R}^{(n-1) \times (n-1)}$

From here take  $Q_2 = \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}$  Then notice that:

$$Q_2^* (Q_1^* A Q_1) Q_2 \quad (36)$$

$$= \begin{bmatrix} I & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & U_2 \end{bmatrix} \quad (37)$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & U_2^* A_1 U_2 \end{bmatrix} \quad (38)$$

then by the same argument as before but this time for  $A_1$

$$= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & A_2 \end{bmatrix} \quad (39)$$

For some new matrix  $A_2$  (this is because we took the eigenvector of  $A_1$ ). We can keep continuing this process until completion. We then have:

$$\left(\prod_i Q_i\right)^* A \prod_i Q_i = \Lambda \quad (40)$$

$$(41)$$

Setting  $Q = \prod_i Q_i$

$$A = Q \Lambda Q^* \quad (42)$$

Where  $\Lambda$  is the diagonal matrix of these eigenvalues.

So it is diagonalizable.

**Exercise 3.4** To prove this take any matrix  $A$ . Note that  $A$  has at least one eigenpair  $\lambda_1, v_1$  because the characteristic polynomial has at least one root. construct  $Q_1 = [v_1, q_2, \dots, q_n]$  where we complete  $v_1$  to an orthonormal basis  $q_2, \dots, q_n$

$$Q_1^* A Q_1 = Q_1 [A v_1 \quad A q_2 \quad \dots \quad A q_n] \quad (43)$$

$$= Q_1 [\lambda_1 v_1 \quad A q_2 \quad \dots \quad A q_n] \quad (44)$$

$$= \begin{bmatrix} \langle \lambda_1 v_1, v_1 \rangle & \langle A q_2, v_1 \rangle & \dots & \langle A q_n, v_1 \rangle \\ \langle \lambda_1 v_1, q_2 \rangle & \langle A q_2, q_2 \rangle & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \langle \lambda_1 v_1, q_n \rangle & \dots & \dots & \langle A q_n, q_n \rangle \end{bmatrix} \quad (45)$$

$$= \begin{bmatrix} \langle \lambda_1 v_1, v_1 \rangle & \langle A q_2, v_1 \rangle & \dots & \langle A q_n, v_1 \rangle \\ 0 & \langle A q_2, q_2 \rangle & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \langle A q_n, q_n \rangle \end{bmatrix} \quad (46)$$

$$(47)$$

Setting  $b = [\langle A q_2, v_1 \rangle, \dots, \langle A q_n, v_1 \rangle]$ :

$$= \begin{bmatrix} \lambda_1 & b \\ 0 & A_1 \end{bmatrix} \quad (48)$$

We can then note that  $A_1$  also has at least one eigenpair by the same argument as before. Call this  $\lambda_2, v_2$ . Complete  $v_2$  to an orthonormal basis of  $\mathbb{R}^{(n-1) \times (n-1)}$ .

$$U_2 = [v_2 \quad z_2 \quad \dots \quad z_{n-1}] \quad (49)$$

$$\text{Set } Q_2 = \begin{bmatrix} I & 0 \\ 0 & U_2 \end{bmatrix}$$

From there note that:

$$Q_2^*(Q_1^*AQ_1)Q_2 = Q_2^* \begin{bmatrix} \lambda & b \\ 0 & A_1 \end{bmatrix} Q_2 \quad (50)$$

$$\begin{bmatrix} I & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & b \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & U_2 \end{bmatrix} \quad (51)$$

$$= \begin{bmatrix} I & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & U_2b \\ 0 & A_1U_1 \end{bmatrix} \quad (52)$$

$$= \begin{bmatrix} \lambda_1 & U_2b \\ 0 & U_2^*A_1U_2 \end{bmatrix} \quad (53)$$

and then by the same argument as before we have that:

$$\begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & A_2 \end{bmatrix} \quad (54)$$

Where \* are just arbitrary numbers that we don't care about. If we continue this process to the end we will be left with an upper triangular matrix we call T.

From here

$$\left(\prod_i Q_i\right)^* A \prod_i Q_i = T \quad (55)$$

$$(56)$$

Setting  $Q = \prod_i Q_i$

$$A = QTQ^* \quad (57)$$

From here note that since eigenvalues are invariant under unitary transformation ( $U^*AU(U^*v) = U^*Av = U^*\lambda v = \lambda U^*v$ , so any eigenvalue  $\lambda$  of A with eigenvector  $v$  is also a eigenvalue of this new matrix with eigenvector  $U^*v$ ). The eigenvalues of A are the eigenvalues of T which we can just read off of the diagonal.

### Exercise 3.5

a) To show this first note that  $\sup |W(A)| = \sup_{x \neq 0} \left| \frac{\langle Ax, x \rangle}{\|x\|^2} \right|$  from here clearly note that:

$$\sup_{x \neq 0} \left| \frac{\langle Ax, x \rangle}{\|x\|^2} \right| \quad (58)$$

$$\geq \left| \frac{\langle Av, v \rangle}{\|v\|^2} \right| \quad (59)$$

for some arbitrary vector v. Choose v to be the eigenvector corresponding to the largest eigenvalue in magnitude.

$$= \left| \frac{\langle Av, v \rangle}{\|v\|^2} \right| = \left| \frac{\langle \lambda_{\max} v, v \rangle}{\|v\|^2} \right| = |\lambda_{\max}| = \rho(A) \quad (60)$$

So the first inequality is true furthermore note that by cuachy shwartz:

$$\sup_{x \neq 0} \left| \frac{\langle Ax, x \rangle}{\|x\|^2} \right| \quad (61)$$

$$\leq \sup_{x \neq 0} \left| \frac{\|Ax\| \|x\|}{\|x\|^2} \right| \quad (62)$$

$$\leq \sup_{x \neq 0} \left| \frac{\|Ax\|}{\|x\|} \right| \quad (63)$$

$$= \|A\|_2 \quad (64)$$

So the second inequality holds as well.

b) so note that  $W$  is a set function so we will view how this operates on arbitrary object in the set. Let  $a \in W(e^{i\theta} A)$  that means  $a = \frac{\langle v, e^{i\theta} Av \rangle}{\|v\|^2}$  for some  $v$ . Which mean:

$$a = \frac{\langle v, e^{i\theta} Av \rangle}{\|v\|^2} \quad (65)$$

$$= e^{i\theta} \frac{\langle v, Av \rangle}{\|v\|^2} \quad (66)$$

Which is  $e^{i\theta} b$  where  $b \in W(A)$  so  $a \in e^{i\theta} W(A)$ . furthermore in the other direction if  $a \in e^{i\theta} W(A)$  then there is a  $v$  such that:

$$= e^{i\theta} \frac{\langle v, Av \rangle}{\|v\|^2} \quad (67)$$

$$= \frac{\langle v, e^{i\theta} Av \rangle}{\|v\|^2} \quad (68)$$

which menas that  $a \in W(e^{i\theta} A)$ . thus  $W(e^{i\theta} A) = e^{i\theta} W(A)$

c) To do this we once again need to consider individual elements. let  $a \in W(A^*)$  then there is a  $v$  such that:

$$a = \frac{\langle v, A^* v \rangle}{\|v\|^2} \quad (69)$$

$$= \frac{\langle Av, v \rangle}{\|v\|^2} \quad (70)$$

$$= \frac{\overline{\langle v, Av \rangle}}{\|v\|^2} \quad (71)$$

$$(72)$$

So thus  $a \in \overline{W(A)} = W(A)^*$  (sorry I am used to the bar meaning conjugate). Conversely let  $a \in W(A)^*$  then there is a  $v$  such that:

$$a = \frac{\langle v, Av \rangle}{\|v\|^2} \quad (73)$$

$$= \frac{\overline{\langle v, Av \rangle}}{\|v\|^2} \quad (74)$$

$$= \frac{\langle Av, v \rangle}{\|v\|^2} \quad (75)$$

$$= \frac{\langle v, A^*v \rangle}{\|v\|^2} \quad (76)$$

$$(77)$$

thus  $a \in W(A^*)$  and as a result  $W(A^*) = W(A)^*$

d) lets take  $a \in W(H(A))$  then there is a  $v$  with:

$$a = \frac{\langle v, \frac{1}{2}(A + A^*)v \rangle}{\|v\|^2} \quad (78)$$

$$= \frac{1}{2} \left( \frac{\langle v, Av \rangle}{\|v\|^2} + \frac{\langle v, A^*v \rangle}{\|v\|^2} \right) \quad (79)$$

$$= \frac{1}{2} \left( \frac{\langle v, Av \rangle}{\|v\|^2} + \frac{\overline{\langle v, Av \rangle}}{\|v\|^2} \right) \quad (80)$$

$$= \text{re} \left( \frac{\langle v, Av \rangle}{\|v\|^2} \right) \quad (81)$$

So this new set is just the real parts of  $W(A)$

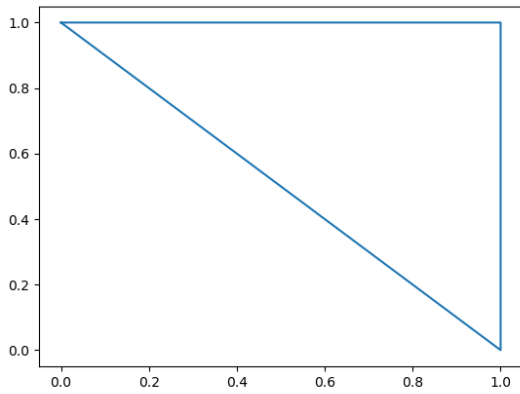
So I think what I would do is since this new matrix is hermitian. its interval should be the distance between the smallest and largest eigenvalues. So as a result I would just use a numerical solver to find both of those eigenvalues and that would give me the range.

### Exercise 3.6

The idea to this one is that we will compute the hermitian part of  $A$  and use that to get the real part of the a point on the numerical range. The eigenvector of this part will correspond to an eigenvector that maps to that point in the numerical range. So I can then plot the Raleigh quotient of those eigenvectors. I then rotate the matrix and repeat.

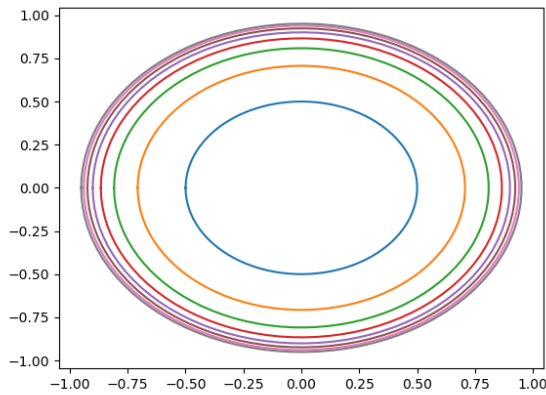
When I trace out the following matrix: I get

$$\begin{bmatrix} 1+i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (82)$$



and everything is included in it.

Here are a bunch of different jordan blocks of increasing dimension:



Thus the numerical range of the jordan normal form seems to be a circle centered at the origin with radius that increases by the size of the dimension. It seems that the radius approaches one but never quite gets there.

So for any arbitrary matrix  $A$ . I would say that the numerical range of the jordan normal form will look like circles of increasing width centered at whatever eigenvalue is degenerate. The width of the cricles is determined by the degeneracy of the eigenvalue. Furthermore taking the convex hull of all of these circles. If the point is not degenerate than the radius is just zero.