HW8

Questions 4.1, 4.2, 4.3, 4.4, 4.5

QUESTION 4.1. *(Easy)* Let A be defined as in equation (4.1). Show that $\det(A) = \prod_{i=1}^{b} \det(A_{ii})$ and then that $\det(A - \lambda I) = \prod_{i=1}^{b} \det(A_{ii} - \lambda I)$. Conclude that the set of eigenvalues of A is the union of the sets of eigenvalues of A_{11} through A_{bb}.

$$A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1b} \\ 0 & A_{22} & \dots & A_{2b} \\ & & & \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_{bb} \end{bmatrix}, \tag{4.1}$$

where each A_{ii} is square.

We have that $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A)\det(D - CA^{-1}B)$

apply schur form to each block and then the whole matrix.

Splitting up A in the following way,

$$A = \left[\begin{array}{c|ccc} A_{11} & A_{12} & \dots & A_{1b} \\ \hline 0 & A_{22} & \dots & A_{2b} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_{bb} \end{array}\right],$$

we can use that

$$\begin{aligned} \det A &= \det A_{11} \cdot \det \left(\begin{bmatrix} A_{22} & \dots & A_{2b} \\ & \ddots & \vdots \\ & & A_{bb} \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} A_{11}^{-1} \begin{bmatrix} A_{12} & \dots & A_{1b} \end{bmatrix} \right) \\ &= \det A_{11} \cdot \det \left(\left[\begin{array}{c|cc} A_{22} & \dots & A_{2b} \\ \hline & \ddots & \vdots \\ & & A_{bb} \end{array}\right] \right). \end{aligned}$$

Partitioning the second matrix as shown and recursing the same logic yields that

$$\det A = \prod_{i=1}^{b} \det A_{ii}$$

As $A - \lambda I$ retains the upper triangular block matrix structure, it follows that

$$\det(A - \lambda I) = \prod_{i=1}^{b} \det(A_{ii} - \lambda I),$$

where each I has the same dimensions as A_{ii} (which is square by assumption).

Denote the set of eigenvalues of A_{ii} as $\lambda(A_{ii})$. Each of these eigenvalues solves the corresponding characteristic equation $\det(A_{ii} - \lambda I) = 0$, meaning that they also solve the characteristic equation of A,

$$\det(A - \lambda I) = \prod_{i=1}^{b} \det(A_{ii} - \lambda I),$$

i.e. that they are also eigenvalues of A.

As each eigenvalue of each block A_{ii} is also an eigenvalue of A,

$$\bigcup_{i=1}^{b} \lambda(A_{ii}) \subseteq \lambda(A).$$

If λ is a solution to the characteristic equation $\det(A - \lambda I) = 0$, then, for at least one i, it must also solve $\det(A_{ii} - \lambda I)$. As a consequence, A does not have any other eigenvalues than those of A_{ii}, and that we have the equality

$$\lambda(A) = \bigcup_{i=1}^{b} \lambda(A_{ii}).$$

QUESTION 4.2. *(Medium; Z. Bai)* Suppose that A is *normal*; i.e., $AA^* = A^*A$. Show that if A is also triangular, it must be diagonal. Use this to show that an n-by-n matrix is normal if and only if it has n orthonormal eigenvectors. Hint: Show that A is normal if and only if its Schur form is normal.

Let A be upper triangular and partitioned as

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

where A_{11} is a scalar, and $A_{12} \in \mathbb{R}^{1\times(n-1)}, \quad A_{22} \in \mathbb{R}^{(n-1)\times(n-1)}$.

$$\begin{aligned} A^*A &= \begin{bmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}^*A_{11} & A_{11}^*A_{12} \\ A_{12}^*A_{11} & A_{12}^*A_{12} + A_{22}^*A_{22} \end{bmatrix} \end{aligned}$$

while

$$\begin{aligned} AA^* &= \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^* & 0 \\ A_{12}^* & A_{22}^* \end{bmatrix} \\ &= \begin{bmatrix} A_{11}A_{11}^* + A_{12}A_{12}^* & A_{12}A_{22}^* \\ A_{22}A_{12}^* & A_{22}A_{22}^* \end{bmatrix} \end{aligned}$$

Since A is normal, these two expressions must be equal, and each block must be equal to each other. Inpecting the first entry,

$$A_{11}A_{11}^* + A_{12}A_{12}^* = A_{11}^*A_{11},$$

meaning that $A_{12}A_{12}^* = 0 \implies A_{12} = 0$ and

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

Continuing the same line of logic on A_{22} yields that if A is triangular and normal, it must be diagonal.

Part 2:
Supposing that A is normal, we can show that Schur form is also normal,

$$\begin{aligned} TT^* &= Q^*AQQ^*A^*Q \\ &= Q^*AA^*Q \\ &= Q^*A^*AQ \\ &= Q^*A^*QQ^*AQ \\ &= T^*T \end{aligned}$$

Since T is both normal and upper triangular, T is also diagonal. Then the eigenvectors of T are the basis unit vectors and form an orthonormal set.

$$\begin{aligned} Te_i &= \lambda_i e_i \\ Q^*AQe_i &= \lambda_i e_i \\ Ax_i &= Q\lambda_i Q^* x_i \\ Ax_i &= \lambda_i x_i, \end{aligned}$$

where the eigenvectors of A, $x_i = Qe_i$, also form an orthonormal basis, since Q is orthogonal.

QUESTION 4.3. *(Easy; Z. Bai)* Let λ and μ be distinct eigenvalues of A, let x be a right eigenvector for λ, and let y be a left eigenvecctor for μ. Show that x and y are orthogonal.

$$\begin{aligned} Ax &= \lambda x \\ y^T A &= \mu y^T \end{aligned}$$

$$\begin{aligned} y^T Ax &= \lambda y^T x \\ (y^T A - \lambda y^T)x &= 0 \\ (\mu y^T - \lambda y^T)x &= 0 \\ (\mu - \lambda)y^T x &= 0 \end{aligned}$$

Either $\mu - \lambda = 0$ (i.e. the eigenvalues are distinct) or $y^T x = 0$ and x and y are orthogonal.

QUESTION 4.4. *(Medium)* Suppose A has distinct eigenvalues. Let $f(z) = \sum_{i=-\infty}^{+\infty} a_i z^i$ be a function which is defined at the eigenvalues of A. Let $Q^*AQ = T$ be the Schur form of A (so Q is unitary and T upper triangular).

1. Show that $f(A) = Qf(T)Q^*$. Thus to compute $f(A)$ it suffices to be able to compute $f(T)$. In the rest of the problem you will derive a simple recurrence formula for $f(T)$.

$$\begin{aligned} f(A) &= \sum_{i=-\infty}^{\infty} a_i A^i \\ &= \dots + a_{-1}A^{-1} + a_0 A^0 + a_1 A^1 + \dots \end{aligned}$$

Let's compute some powers of the Schur form and see what we get,

$$\begin{aligned} (Q^*AQ)^2 &= Q^*A^2Q \\ (Q^*AQ)^3 &= Q^*A^3Q \\ (Q^*AQ)^{-1} &= Q^*A^{-1}Q \\ (Q^*AQ)^{-2} &= (Q^*A^{-1}Q)^2 = Q^*A^{-2}Q. \end{aligned}$$

It is evident that we can write $f(T)$ as

$$\begin{aligned} f(T) &= \sum_{i=-\infty}^{\infty} a_i (Q^*AQ)^i \\ &= \sum_{i=-\infty}^{\infty} a_i Q^*A^iQ \\ &= Q^* \left(\sum_{i=-\infty}^{\infty} a_i A^i \right) Q \\ f(T) &= Q^* f(A) Q \\ Qf(T)Q^* &= f(A) \end{aligned}$$

which is what we wanted to prove.

2. Show that $(f(T))_{ii} = f(T_{ii})$ so that the diagonal of $f(T)$ can be computed from the diagonal of T.

$$\begin{aligned} (f(T))_{ii} &= \left(\sum_{j=-\infty}^{\infty} a_j T^j \right)_{ii} \\ &= \dots + (a_{-1}T^{-1})_{ii} + (a_0 T^0)_{ii} + (a_1 T^1)_{ii} + \dots \end{aligned}$$

We can write $T = D(X + I)$, where D is diagonal and X is strictly upper triangular. Then,

$$\begin{aligned} T &= D(X+I) \\ T^k &= D^k (X+I)^k \\ &= D^k (\text{strictly upper triangular matrices} + I^k), \end{aligned} \tag{1}$$

where I have used that X^k is strictly upper triangular (HW1). We see that the diagonal of T^k is its diagonal elements to the k-th power, D^k.

Further, the inverse of T is

$$T^{-1} = D^{-1}(X+I)^{-1} = D^{-1}(I - X + X^2 + \cdots + (-1)^n X^n).$$

We see that it's diagonal entries is D^{-1}, as the rest of the matrices are strictly upper triangular. Generalising to an arbitrary inverse power,

$$T^{-k} = (T^{-1})^k,$$

we observe that we could apply the same logic as in (1), and get that the diagonal entries of T^{-k} are the diagonal entries of T^{-1} raised to the k-th power. In conclusion,

$$(T^k)_{ii} = (T_{ii})^k$$

for all integers k, and therefore

$$\begin{aligned}
(f(T))_{ii} &= \cdots + (a_{-1}T^{-1})_{ii} + (a_0 T^0)_{ii} + (a_1 T^1)_{ii} + \ldots \\
(f(T))_{ii} &= \cdots + a_{-1}(T_{ii})^{-1} + a_0 (T_{ii})^0 + a_1 (T_{ii})^1 + \ldots \\
(f(T))_{ii} &= f(T_{ii}),
\end{aligned}$$

and since f is defined at the eigenvalues of A, which appear in the diagonal of T, this is well-defined.

3. Show that $Tf(T) = f(T)T$.

$$\begin{aligned}
Tf(T) &= T \sum_{i=-\infty}^{\infty} a_i T^i \\
&= \sum_{i=-\infty}^{\infty} a_i T^{i+1} \\
&= \sum_{i=-\infty}^{\infty} a_i T^i T \\
&= f(T)T
\end{aligned}$$

4. From the last result, show that the ith superdiagonal of $f(T)$ can be computed from the $(i-1)$st and earlier subdiagonals. Thus, starting at the diagonal of $f(T)$, we can compute the first superdiagonal, second superdiagonal, and so on.

Errata: Page 188, Question 4.4, Part 4: "earlier subdiagonals" should be "earlier superdiagonals". (Yulong Dong)

We note from 4.3.2 that $f(T)$ is also upper triangular. So far, we know the diagonal elements,

$$Tf(T) = f(T)T$$

$$\begin{bmatrix} T_{11} & \ldots & T_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & T_{nn} \end{bmatrix} \begin{bmatrix} f(T_{11}) & * & * \\ 0 & \ddots & * \\ 0 & 0 & f(T_{nn}) \end{bmatrix} = \begin{bmatrix} f(T_{11}) & * & * \\ 0 & \ddots & * \\ 0 & 0 & f(T_{nn}) \end{bmatrix} \begin{bmatrix} T_{11} & \ldots & T_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & T_{nn} \end{bmatrix}$$

The equality must holde index-wise,

$$(Tf(T))_{i,j} = \sum_{k=1}^{n} T_{i,k}(f(T))_{k,j} = \sum_{k=1}^{n}(f(T))_{ik}T_{kj} = (f(T)T)_{ij}$$

As both matrices are upper triangular, $T_{ij} = 0$ and $f(T)_{ij} = 0$ for $i > j$, and this equality becomes

$$\sum_{k=i}^{j} T_{ik}(f(T))_{kj} = \sum_{k=i}^{j}(f(T))_{ik}T_{kj}.$$

For notational simplicity, let $F_{ij} = (f(T))_{ij}$, and the sum is

$$\sum_{k=i}^{j} T_{ik}F_{kj} = \sum_{k=i}^{j} F_{ik}T_{kj}.$$

First, we can let $j = i + 1$, and

$$\begin{aligned} T_{i,i}F_{i,i+1} + T_{i,i+1}F_{i+1,i+1} &= F_{i,i}T_{i,i+1} + F_{i,i+1}T_{i+1,i+1} \\ F_{i,i+1} &= T_{i,i+1}\frac{F_{ii} - F_{i+1,i+1}}{T_{i,i} - T_{i+1,i+1}} \end{aligned}$$

Note that the divisor is non-zero as the diagonal elements of T are the eigenvalues of A, and the eigenvalues are distinct by assumption.
The first superdiagonal of F is well-defined and can be calculated from the diagonal.
In the sum above, we can let $j = i + 2$ and

$$T_{ii}F_{i,i+2} + T_{i,i+1}F_{i+1,i+2} + T_{i,i+2}F_{i+2,i+2} = F_{i,i}T_{i,i+2} + F_{i,i+1}T_{i+1,i+2} + F_{i,i+2}T_{i+2,i+2}.$$

$$F_{i,i+2} = \frac{(F_{i,i} - F_{i+2,i+2})T_{i,i+2} + (F_{i,i+1} - F_{i+1,i+2})T_{i+1,i+2}}{T_{i,i} - T_{i+2,i+2}}$$

Like this we can compute the second super diagonal from the first superdiagonal and the diagonal.

To find the k-th superdiagonal, we can let $j = i + k$, which results in the equation

$$F_{i,i+k} = \frac{1}{T_{i,i} - T_{i+k,i+k}} \sum_{p=i}^{i+k-1} (F_{i,p} - F_{p,i+k})T_{p,i+k},$$

which involves all earlier superdiagonals as well as the diagonal.

QUESTION 4.5. *(Easy)* Let A be a square matrix. Apply either Question 4.4 to the Schur form of A or equation (4.6) to the Jordan form of A to conclude that the eigenvalues of $f(A)$ are $f(\lambda_i)$, where the λ_i are the eigenvalues of A. This result is called the *spectral mapping theorem.*

This question is used in the proof of Theorem 6.5 and section 6.5.6.