APPLIED LINEAR ALGEBRA HOMEWORK 2

TOMMENIX YU
STAT 31430
DUE FRIDAY, OCT. 21, 3PM

1. Written Assignment

Exercise 1.1. *Suppose that $A \in \mathcal{M}_{n,n}(\mathbb{C})$ has eigenvalue $\lambda \in \mathbb{C}$.We say that $v \in \mathbb{C}^n \backslash \{0\}$ is a generalized eigenvector of degree $k \geq 1$, associated to the eigenvalue λ, if $x \in \ker((A-\lambda I)^k)$ but $x \notin \ker((A-\lambda I)^{k-1})$.*

(a) Show that if $v \in \mathbb{C}^n \backslash \{0\}$is a generalized eigenvector of degree k for some $k \geq 1$, then

$$\{v, (A-\lambda I)v, (A-\lambda I)^2 v, \ldots, (A-\lambda I)^{k-1} v\}$$

is a linearly independent set of vectors in $\mathbb{C}^n$.

(b) Fix $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1 \neq \lambda_2$. Let $u_1, \ldots, u_5 \in \mathbb{R}^5$ be a collection of five linearly independent vectors in $\mathbb{R}^5$, and let U be the 5×5 matrix with columns given by

$$U = \left[u_1, u_2, u_3, u_4, u_5\right].$$

Set

$$A = UJU^{-1}, \quad J = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix},$$

$$J_1 = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{bmatrix}, \quad J_2 = \begin{bmatrix} \lambda_2 & 1 \\ 0 & \lambda_2 \end{bmatrix}$$

Show that u_1 is an eigenvector of A, that u_2 is a generalized eigenvector of degree 2, and that u_3 is a generalized eigenvector of degree 3. Similarly, show that u_4 is an eigenvector of A, and u_5 is a generalized eigenvector of degree 2. What are the associated eigenvalues in each case?

Remark: This is an example of an expression of the matrix A in Jordan canonical form, a generalization of the idea of diagonalizing a matrix. It can be shown that every matrix $A \in \mathcal{M}_{n,n}(\mathbb{C})$ can be reduced to Jordan canonical form by a suitable change of basis.

Proof.

(a): We do it by contradiction. Assume that the set is not linearly independent, then there exist non-zero $\tilde{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_{k-2})$ such that

$$(A - \lambda I)^{k-1} v = \sum_{i=0}^{k-2} \alpha_i (A - \lambda I)^i v$$

multiplying both side with $(A - \lambda I)$ on the left yields

$$\sum_{i=1}^{k-1} \alpha_{i-1} (A - \lambda I)^i v = (A - \lambda I)^k v = 0$$

which means that

$$S_1 := \{(A - \lambda I)v, (A - \lambda I)^2 v, \dots, (A - \lambda I)^{k-1} v\}$$

is not linearly independent (since unique expression of 0 under any basis).

We can repeat the above process on S_1 (with suitable non-zero $\tilde{\beta} = (\beta_1, \dots, \beta_{k-2})$) and we can show that

$$S_2 := \{(A - \lambda I)^2 v, \dots, (A - \lambda I)^{k-1} v\}$$

is not linearly independent.

Similarly, we can show that for any $n = 1, 2, \dots, k-2$,

$$S_j := \{(A - \lambda I)^j v, \dots, (A - \lambda I)^{k-1} v\}$$

is not linearly independent. In particular, when $j = k - 2$, we have

$$(A - \lambda I)^{k-2} v = c(A - \lambda I)^{k-1} v$$

now we multiple both side with $(A - \lambda I)$ on the left one last time to get

$$(A - \lambda I)^{k-1} v = c(A - \lambda I)^k v = 0$$

contradiction to the assumption that v is degree k. Thus the set in problem is linearly independent.

(b):

Lemma 1.1. *For any $n \in \mathbb{Z}$, let $N_n^1 = N^1 \in \mathcal{M}_n(\mathbb{C})$ be the matrix*

$$N^1 = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & 0 \end{bmatrix}$$

i.e. (the ith row, jth column of N^1) $n_{i,j}^1 = \delta_{i,j-1}$, where δ is the Kronecker delta.

Similarily, define $N^k = \left(n_{i,j}^k\right)_{i,j}$ where $n_{i,j}^k = \delta_{i,j-k}$. Then for any $m, l \in \mathbb{Z}^*$ such that $m + l \leq n$, $N^m N^l = N^{m+l}$.

Proof. (of Lemma 1.1) The proof is simply by computation.

Let $N^m N^l = \left(n'_{i,j}\right)_{i,j}$, then

$$n'_{i,i+(m+l)} = \sum_{k=1}^{n} n^m_{i,k} n^l_{k,i+(m+l)} = 0 + \cdots + 0 + n^m_{i,i+m} n^l_{i+m,i+(m+l)} + 0 + \cdots + 0 = 1$$

and for $j \neq i + (m + l)$,

$$n'_{i,j} = \sum_{k=1}^{n} n^m_{i,k} n^l_{k,j} = 0 + \cdots + 0 + n^m_{i,i+m} \cdot 0 + \cdots + 0 \cdot n^l_{i+m,j-l} + 0 + \cdots + 0 = 0$$

which means $n'_{i,j} = \delta_{i,j-(m+l)} = n^{m+l}_{i,j}$, and we are done. □

Now, note that $(J - \lambda_1 I) = \begin{bmatrix} N_3^1 & 0 \\ 0 & B \end{bmatrix} \Rightarrow (J - \lambda_1 I)^m = \begin{bmatrix} N_3^m & 0 \\ 0 & B^m \end{bmatrix}$ for $m \leq 3$ by Lemma 1.1. Therefore

$$(A - \lambda_1 I)^m = (UJU^{-1} - U\lambda_1 IU^{-1})^m = U(J - \lambda_1 I)^m U^{-1}$$

since $(A-\lambda_1 I)^m u_m$ is the m-th column of $(A-\lambda_1 I)^m U = U(J-\lambda_1 I)^m U^{-1} U = U(J-\lambda_1 I)^m$, by definition the m-th column of N_3^m is 0, so is the m-th column of $U(J - \lambda_1 I)^m$.

Again, by definition of N_3^m, the m-th column of N_3^{m-1} is non-zero, and the m-th column of N_3^{m+1}, and since U unitary, u_m is a degree m eigenvector for $1 \leq m \leq 3$.

For $m = 4, 5$, the computation is exactly the same (except this time N is the second diagonal block), and we get the wanted result.

(This proof is not as direct as just doing every computation for the 5x5 matrix. However, this can be generated for all matrices, thus a proof for the Jordan reduction (with change of basis added).)

□

Exercise 1.2.

(a) Show that for every $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $\mathrm{Im}(A) \cap \ker(A^*) = 0$. (Hint: For $x \in \mathrm{Im}(A) \cap \ker(A^*)$, write $x = Ay$ and look at $< x, x >$.)

(b) For $A \in \mathcal{M}_{n,n}(\mathbb{C})$ with entries given by $A = (a_{ij})_{i,j=1}^n$, the Frobenius norm is defined via

$$||A||_F^2 = \sum_{i,j=1}^n |a_{ij}|^2.$$

Show that for $A \in \mathcal{M}_{n,n}(\mathbb{C})$, $||A||_F^2 = \mathrm{tr}(AA^*)$ and $||A||_F^2 = ||UAU^*||_F^2$ for all unitary U.

(c) Let $||\cdot||_F^2$ be as in (b) above. Suppose that $A \in \mathcal{M}_{n,n}(\mathbb{C})$ has eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ (possibly repeated). Show that A is normal if and only if $||A||_F^2 = \sum_{i=1}^n |\lambda_i|^2$.

Proof.

(a): Let $x \in \mathrm{Im}(A) \cap \ker(A^*)$, write $x = Ay$ since it's in the image, then since it's in the kernel we can write $A^*Ay = 0$. If $y = 0$ then we are done since then $x = 0$, so we only prove that when $A^*A = 0$, $A = 0$ such that $\mathrm{Im}(A) = 0$, which yields the result.

By Schur Decomposition we get $A = UTU^*$ and thus $A^* = \overline{(UTU^*)^{-1}} = U^*T^*U$, so we have $A^*A = 0 \Rightarrow U^*T^*TU = 0 \Rightarrow T^*T = 0$. We can choose T upper triangular and as we have seen in class, each entry of T^*T is the some of many norm squares of the entries of T, so each ≥ 0. So the only possibility for $T^*T = 0$ is for each term in T be 0, which implies $T = 0$, which implies $A = 0$, and we are done.

(b):

$$\mathrm{tr}(AA^*) = \sum_{i=1}^n \sum_{j=1}^n |a_{ji}|^2 = \sum_{i,j=1}^n |a_{ij}|^2$$

by direct computation.

Lemma 1.2. *The trace of a matrix is the sum of its eigenvalues counted with multiplicity. The determinant is the product of all eigenvalues.*

Proof. (Lemma 1.2)

Trace:

The $n-1$ degree term in P_A has coefficient $(-1)^n(\mathrm{tr}(A))$ since in the computation of determinant, in order to get λ^{n-1}, we have to choose $n-1$ element on the diagonal, but the last multiplier can only be the constant part of the term left on the diagonal.

But on the other hand, $P_A = (-1)^n(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$, which has $n-1$ degree term's coefficient as the sum of the eigenvalues. Combined, this means that the trace is the sum of the eigenvalues. Thus we're done.

Determinant:

This is similar, note that P_A has 0 degree term equals to the summand where no λ is chosen in the definition of determinant, the 0 degree term is the determinant of the matrix.

By from $P_A = (-1)^n(\lambda-\lambda_1)\cdots(\lambda-\lambda_n)$ this expression it is also clear that the 0 degree term is nothing but the product of all eigenvalues, where the negative signs cancel out perfectly. □

We know that $||A||_F^2 = \mathrm{tr}(AA^*)$ and $||UAU^*||_F^2 = \mathrm{tr}(UAA^*U^*)$. Further, B and UBU^* has the same eigenvalues since they can be reduced to the same Schur decomposition, and the diagonal of the triangular matrix are the eigenvalues.

Thus, by Lemma 1.2

$$\begin{aligned} ||A||_F^2 &= \mathrm{tr}(AA^*) = \text{ sum of eigenvalues of } AA^* \\ &= \text{ sum of eigenvalues of } UAA^*U^* = \mathrm{tr}(UAA^*U^*) = ||UAU^*||_F^2. \end{aligned}$$

(c): ($\Rightarrow$): Suppose A is normal, then $\exists U$ unitary such that $A = UDU^*$ where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then

$$||A||_F^2 = ||UAU^*||_F^2 = \sum_{i=1}^{n} |\lambda_i|^2.$$

($\Leftarrow$): By Schur decomposition,

$$\sum_{i=1}^{n} |\lambda_i|^2 = ||A||_F^2 = ||U^*TU||_F^2 = ||T||_F^2 = \sum_{i=1}^{n} |\lambda_i|^2 + \sum_{i\neq j}^{n} |t_{ij}|^2$$

$$\Rightarrow \sum_{i\neq j}^{n} |t_{ij}|^2 = 0 \Rightarrow T \text{ is diagonal.}$$

But then A is unitarily diagonalizable, which means A is normal by theorem in class. □

Exercise 1.3.

(a) Show that for every invertible upper triangular matrix $T \in \mathcal{M}_n(\mathbb{C})$ there exists $S \in \mathcal{M}_n(\mathbb{C})$ such that $S^2 = T$.

(b) Show that for every invertible matrix $A \in \mathcal{M}_n(\mathbb{C})$ there exists $B \in \mathcal{M}_n(\mathbb{C})$ such that $B^2 = A$.(Hint: Use the Schur decomposition to reduce to (a).)

(c) We say that a Hermitian matrix $A \in \mathcal{M}_n(\mathbb{C})$ is positive definite if every eigenvalue $\lambda \in \sigma(A)$ satisfies $\lambda > 0$. Show that if $A \in \mathcal{M}_n(\mathbb{C})$ is a Hermitian positive definite matrix, then there exists a unique Hermitian positive definite $B \in \mathcal{M}_n(\mathbb{C})$ such that $B^2 = A$. (Hint: One can either elaborate on (a)–(b) above, or, essentially equivalently, diagonalize A.)

Proof.

(a): T is invertible means that $\det(T) \neq 0$ and since $\det(T) = \prod_{i=1}^{n} t_{ii}^2$, non of the diagonal of T is 0.

I will prove the existence of 1 particular upper triangular S (since the square root of T might not even be triangular) such that $S^2 = T$. I prove this by induction (though we want to stop at the n-th column) on the existence of columns of S.

For $i = 1$, we want to settle down the first column of S. Since we want an upper triangular S, we let anything below the diagonal be 0. As for s_{11}, we know that it is a square root of $t_{11} \neq 0$. If $t_{11} \in \mathbb{R}^+$, we choose $s_1 1 = \sqrt{t_{11}}$; otherwise we choose the complex square root of t_{11} such that $\text{Im}(s_{11}) > 0$. This choice is unique and guaranteed to exist.

IH says that we have chosen (in the same way, especially the diagonal entries) the first $N - 1$ columns, then we want to choose the N-th column. Still, we first let all entries below the diagonal 0. Then, for the diagonal term we still choose the square root of t_{NN} that is either positive real or has positive imaginary part.

Then, for $s_{N-1,N}$, it is the solution of the equation

$$s_{N-1,N-1} \cdot s_{N-1,N} + s_{N-1,N} \cdot s_{N,N} = t_{N-1,N}$$

which is guaranteed to have a unique solution since $s_{N,N} + s_{N-1,N-1} \neq 0$, exactly because the way we choose their value in the complex plane.

We will then see that each entry above in this column can be written as $\frac{a}{b}$ where a is the corresponding entry in T minus some terms in S, and b is the sum of some diagonal entries in S, thus not 0.

So we've proven that at least some S exists such that $S^2 = T$.

(b): Since $A = UTU^*$ where $T = S^2$. Let $B = USU^*$ then $B^2 = A$.

(c): Since A is positive definite, it is invertible (no 0 eigenvalue), which means there exists $B = USU^*$. We've seen in part (b) $B^2 = A$.

Now we will show B is positive definite, Hermitian, and unique.

Positive definite: Since A is positive definite, so T has positive diagonal. By our choice of the diagonal of S in (a), S also has positive diagonal, which means $B = USU^*$ is positive definite. In addition, it means that B is invertible.

Hermitian:

$$B^2 = A = A^* = (B^2)^* = \overline{(BB)^{-1}} = \overline{B^{-1}B^{-1}} = (B^*)^2$$

which implies

$$US^2U^* = B^2 = (B^*)^2 = U\left(S^*\right)^2 U^* \quad \Rightarrow \quad S^2 = (S^*)^2 \quad \Rightarrow \quad T = T^*$$

But in class we've discussed (in the proof that if A is normal it is unitarily diagonalizable) that if a triangular T is normal, then it is diagonal. Applying it (Hermitian ⇒ normal) we see that T is diagonal here, and so is S due to our construction in (a). Thus $S^2 = (S^*)^2 \Rightarrow S = S^* \Rightarrow B = B^*$, so B is Hermitian.

Unique: Suppose there are two positive definite Hermitian matrix B, C such that $B^2 = C^2 = A$. We want to show that $B = C$.

Notice that $B^2 = US^2U^*$ where S^2 is diagonal and has the exactly same diagonal entries as the eigenvalues of A. Similar for $C^2 = PD^2P^*$, so $S^2 = D^2$ and since they are diagonal $S = D$. So $US^2U^* = PS^2P^*$ means that P and U is the same change of variable matrix, so they are the same.

Thus, $B = C$ and it is unique.

□

Exercise 1.4. *Recall that for $S \in \mathcal{M}_n(\mathbb{C})$ and $\lambda \in \sigma(S)$, the eigenspace associated to λ is*

$$E_\lambda^{(S)} := \{x \in \mathbb{C} | Ax = \lambda x\}.$$

(1) Show that if $S, T \in \mathcal{M}_n(\mathbb{C})$ are such that $ST = TS$ then for all $\lambda \in \sigma(S)$ we have

$$T(E_\lambda^{(S)}) \subset E_\lambda^{(S)}.$$

(2) Show that for all $T \in \mathcal{M}_n(\mathbb{C})$ if $S \in \mathcal{M}_n(\mathbb{C})$ is Hermitian and

$$T(E_\lambda^{(S)}) \subset E_\lambda^{(S)} \quad \textit{for all} \quad \lambda \in \sigma(S)$$

then $TS = ST$.

(In fact, a bit more can be said: a sample result of this type is that if $S, T \in \mathcal{M}_n(\mathbb{C})$ are diagonalizable, then $ST = TS$ if and only if S and T can be diagonalized simultaneously, i.e. there exists $P \in \mathcal{M}_n(\mathbb{C})$ invertible such that $P^{-1}SP$ and $P^{-1}TP$ are both diagonal. You don't have to prove this for this exercise.)

Proof.

(a): $\forall v \in E_\lambda^{(S)}$,

$$Sv = \lambda v \quad \Rightarrow \quad S(Tv) = TSv = T\lambda v = \lambda(Tv)$$

which means that $Tv \in E_\lambda^{(S)}$.

(b): Since S is Hermitian, we know that $\mathbb{C}^n = \bigoplus E_{\lambda_i}^{(s)}$, which means that $\forall v \in \mathbb{C}^n$, we can write v as a linear combination of eigenvectors of S. Further, since for any eigenvector u_i associated to λ_i, we have

$$S(Tu_i) = \lambda T u_i = T\lambda u_i = TSu_i.$$

Thus ST is the same linear map as TS since they map all vectors in a basis to the same output vector, and that the basis is complete.

□

Exercise 1.5. *Suppose that $A, B \in \mathcal{M}_n(\mathbb{C})$ are two orthogonal matrices.*

(1) Find $C \in \mathcal{M}_n(\mathbb{C})$such that $A + B = ACB$
(2) Show that if $\det(A) + \det(B) = 0$, then $\det(A + B) = 0$. (Hint: You may find it useful to use part (a). A first step from there is to see what you can say about the value of the product $\det(A)\det(B)$ in this situation.)

Proof.

(a): Since A, B are orthogonal, their inverse exists. So $C := A^{-1} + B^{-1}$ is well defined. And it is indeed the matrix we want since

$$ACB = A(A^{-1} + B^{-1})B = B + A = A + B.$$

(b): By Lemma 1.2, we know that the determinant is the product of all eigenvalues and thus A has the same determinant with A^T, and A orthogonal implies $A^T = A^{-1}$.

Now note that

$$\det(A)^2 = \det(A)\det(A^{-1}) = \det(I) = 1$$

we see $\det(A) = \pm 1$. Same applies to B.

If $\det(A) + \det(B) = 0$ then WLOG we assume $\det(A) = 1, \det(B) = -1$.

Then

$$\begin{aligned}\det(A + B) &= \det(ACB) = \det(A)\det(A^{-1} + B^{-1})\det(B) \\ &= -\det(A^T + B^T) = -\det((A + B)^T) \\ \text{(again by lemma 1.2)} &= -\det(A + B)\end{aligned}$$

which implies $\det(A + B) = 0$.

$\square$

Exercise 1.6. *Let $A \in \mathcal{M}_n(\mathbb{C})$ be a given Hermitian positive definite matrix.*

(1) Show that for all $B \in \mathcal{M}_n(\mathbb{C})$ if B is Hermitian, than there exists an invertible matrix $P \in \mathcal{M}_n(\mathbb{C})$ such that both

$$P^*AP = I \quad \text{and} \quad P^*BP \ \text{ is diagonal}$$

hold simultaneously. (Hint: You may find it useful to use the result of 3(c) above.)

(2) We say that a Hermitian matrix $A \in \mathcal{M}_n(\mathbb{C})$ is positive semi-definite if every $\lambda \in \sigma(A)$ satisfies $\lambda \geq 0$. Show that if $A \in \mathcal{M}_n(\mathbb{C})$ is Hermitian positive semi-definite, then $\langle Bx, x\rangle \geq 0$ for all $x \in \mathbb{C}^n$.

(3) Use parts (a) and (b) above to show that if B is a Hermitian positive semi-definite matrix in $M_N(\mathbb{C})$, then $\det(A) \leq \det(A+B)$.

Proof.

(a): By 3(c), $A = CC^*$ where C is positive definite Hermitian (since $C^2 = CC^*$). Then, we have

$$C^{-1}A(C^*)^{-1} = I$$

where the inverses exist as C is positive definite.

Then $C^{-1}B(C^*)^{-1}$ is Hermitian since

$$(C^{-1}B(C^*)^{-1})^* = ((C^*)^{-1})^*B^*(C^{-1})^* = C^{-1}B(C^*)^{-1}$$

and thus it can be unitarily diagonalized, so

$$C^{-1}B(C^*)^{-1} = UDU^*.$$

Let $P = (C^*)^{-1}U$, then

$$\begin{cases} P^*AP = U^*IU = I \\ P^*BP = D \end{cases}$$

thus we are done.

(b): Since A is Hermitian, $\langle Ax, x\rangle = \langle x, Ax\rangle$, and it is usually the definition of semi-definite matrix to say that $\langle x, Ax\rangle \geq 0$. However, we define it here differently so I'd rather prove it as a property.

Lemma 1.3. *For A semi-positive definite (Hermitian by definition), $\langle x, Ax\rangle \geq 0$ for all x.*

Proof. (Lemma 1.3)

Any x can be written as a linear combination of eigenvectors of A since A Hermitian means $\mathbb{C}^n = \bigoplus E_{\lambda_i}$. To be explicit let's just write out $x = \sum_{i=1}^n c_i v_i$ for v_i linearly independent (just explained existence).

But

$$\langle c_i v_i, Ac_i v_i\rangle = \langle c_i v_i, \lambda_i c_i v_i\rangle = \lambda_i |c_i|^2 ||v_i||_2^2 \geq 0$$

and thus

$$\langle x, Ax \rangle = \sum_{i=1}^{n} \lambda_i |c_i|^2 ||v_i||_2^2 \geq 0.$$

$\square$

Lemma 1.3 basically tells the exact same point as the required: $\langle Bx, x \rangle \geq 0$ for all $x \in \mathbb{C}^n$ for semi-positive definite B.

(c): Let P be defined as in (a). Then we have

$$\begin{cases} A = (P^*)^{-1} I P^{-1} \\ B = (P^*)^{-1} D P^{-1} \end{cases} \Rightarrow (A + B) = (P^*)^{-1}(I + D)P^{-1}$$

Also, since $A = (P^*)^{-1} I P^{-1}$ we know $\det(A) = \det(P^*)^{-1} \det(P^{-1})$.

One last piece to the puzzle is that D is a diagonal matrix with each diagonal entries ≥ 0, which implies $(I + D)$ is diagonal and has each diagonal entries ≥ 1 (it's real because B Hermitian).

Thus

$$\det(A + B) = \det(P^*)^{-1} \det(I + D) \det(P^{-1}) = \det(A) \det(I + D)$$

where $\det(I + D)$ is the product of its diagonal entries, which is ≥ 1 since each one is, so

$$\det(A + B) \geq \det(A).$$

$\square$