MATH36001 Solutions: Theory of Eigensystems 2015

Note: the solution to Example 5 in the handout is added at the end of this document.

1. The characteristic polynomial of A is given by $p(t) = \det(tI - A) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$ so that $p(0) = (-1)^n \det(A) = (-1)^n \prod_{i=1}^n \lambda_i$.

One can show by induction that the coefficient of t^{n-1} in $p(t) = \det(tI - A)$ is $-\operatorname{trace}(A)$ and that of $p(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$ is $-\sum_{i=1}^n \lambda_i$.

(*A direct proof is obtained by using the Schur form* $A = UTU^*$, *and noting that* $\operatorname{trace}(A) = \operatorname{trace}((UT)U^*) = \operatorname{trace}(U^*UT) = \operatorname{trace}(T)$, *where we have used the relation* $\operatorname{trace}(AB) = \operatorname{trace}(BA)$ *proven in exercise 9 (ii) of the basics handout.*)

Since the eigenvalues of A^k are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$, it follows that $\operatorname{trace}(A^k) = \sum_{i=1}^n \lambda_i^k$.

2. (a) Let $x = \sum_{i=1}^k \alpha_i x_i \in \mathcal{S}$. Then $Ax = \sum_{i=1}^k \alpha_i \lambda_i x_i \in \mathcal{S}$ so that $\mathcal{S}$ is an invariant subspace under A.

(b) We have $Ae_1 = 2e_1 \in \mathcal{S}$ and $Ae_2 = e_1 + 2e_2 \in \mathcal{S}$ so $\mathcal{S}$ is an invariant subspace under A. Note that e_2 is not an eigenvector.

3. Let $U_k = [\, u_1 \quad \ldots \quad u_k \,]$, e_i is the ith column of the identity matrix and T_k denotes the $k \times k$ leading principal submatrix of T. Rewrite $A = UTU^*$ as $AU = UT$ so that

$$AU_k = AU\,[\, e_1 \quad \ldots \quad e_k \,] = UT\,[\, e_1 \quad \ldots \quad e_k \,] = U \begin{bmatrix} T_k \\ O \end{bmatrix} = U_k T_k$$

and by Theorem 1, U_k is invariant under A.

4. (a) If $A^*A = AA^*$ then $B^*B = U^*A^*UU^*AU = U^*A^*AU = U^*AA^*U = U^*AUU^*A^*U = BB^*$ so B is normal. If $A^* = A$ then $B^* = U^*A^*U = U^*AU = B$ so B is Hermitian.

(b) Take $S = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ and $A = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$ which is Hermitian and therefore normal. Then $S^{-1}AS = \begin{bmatrix} 2 & 0 \\ 2 & -2 \end{bmatrix}$ is neither Hermitian nor normal.

5. Hermitian, skew-Hermitian and unitary matrices are normal. If A is normal, there exist a unitary matrix U and a diagonal matrix $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ such that $A = U\Lambda U^*$.

(a) A Hermitian $\Rightarrow A = U\Lambda U^* = (U\Lambda U^*)^* = U\overline{\Lambda}U^* \Rightarrow \Lambda = \overline{\Lambda}$, i.e., $\Lambda \in \mathbb{R}^{n\times n}$. Conversely, A normal with real eigenvalues $\Rightarrow A^* = (U\Lambda U^*)^* = U\Lambda U^* = A$ so A is Hermitian.

(b) Proof similar to (a). Here, $\Lambda = -\overline{\Lambda}$, i.e., $\Lambda \in i\mathbb{R}^{n\times n}$.

(c) A unitary $\Rightarrow I = A^*A = U\overline{\Lambda}U^*U\Lambda U^* = U\overline{\Lambda}\Lambda U^* \Rightarrow \overline{\Lambda}\Lambda = I$, i.e., $|\lambda_i| = 1$, $i = 1, \ldots, n$. Conversely, A normal with eigenvalues on the unit circle $\Rightarrow A^*A = U\overline{\Lambda}U^*U\Lambda U^* = I \Rightarrow A$ is unitary.

6. (a) If N is nilpotent then $N^k = O$ for some $k > 0$. Let λ be an eigenvalue of N with corresponding eigenvector x. Then $N^2x = \lambda^2 x$, $N^3 x = \lambda^3 x, \dots, N^k x = \lambda^k x = 0$. Hence $\lambda = 0$ since $x \neq 0$.

(b) Consider the generic 4×4 strictly upper triangular matrix $T = \begin{bmatrix} 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 \end{bmatrix}$, where $\times$ denotes any entry that is not necessarily zero. Then,

$$T^2 = \begin{bmatrix} 0 & 0 & \times & \times \\ 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad T^3 = \begin{bmatrix} 0 & 0 & 0 & \times \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad T^4 = O.$$

In general, powering T causes the nonzero superdiagonals to move a diagonal at a time towards the top right-hand corner until there is none left. If T is $n \times n$ then necessarily $T^n = O$.

(c) Let $N = UTU^*$ be the Schur decomposition of N. If all eigenvalues of N are zero then T is strictly upper triangular and therefore nilpotent. Since $N^k = UT^kU^*$ it follows that if T is nilpotent so is N.

7. When A is normal the spectral theorem says that there exist a unitary matrix $U = [\, u_1 \quad \dots \quad u_n \,]$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ such that

$$A = U\Lambda U^* = [\, \lambda_1 u_1 \quad \dots \quad \lambda_n u_n \,] \begin{bmatrix} u_1^* \\ \vdots \\ u_n^* \end{bmatrix} = \sum_{i=1}^{n} \lambda_i u_i u_i^*.$$

A has two eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 3$ with corresponding eigenvectors $x_1 = (1/\sqrt{2})\,[-1 \quad 1]^T$ and $x_2 = (1/\sqrt{2})\,[1 \quad 1]^T$. Therefore

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} [-1 \quad 1] + \frac{3}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} [1 \quad 1].$$

8. Let T be a normal upper triangular matrix. That T is diagonal is seen by equating the entries of T^*T and TT^* as follows.

$e_1^T(T^*T)e_1 = e_1^T TT^* e_1 \;\Leftrightarrow\; t_{11}\bar{t}_{11} = t_{11}\bar{t}_{11} + \sum_{j=2}^{n} t_{1j}\bar{t}_{1j}$ which implies $0 = \sum_{j=2}^{n} |t_{1j}|^2$, i.e., $t_{1j} = 0$ for $j = 2, \dots, n$. In a similar way we show that $e_2^T(T^*T)e_2 = e_2^T TT^* e_2$ implies that $t_{2j} = 0$ for $j = 3, \dots, n$. Arguing for each diagonal entry in turn, we conclude that $t_{ij} = 0$, $j > i$, $i = 1, \dots, n$. Since T is upper triangular we have shown that T is diagonal.

9. (a) 2 Jordan blocks so 2 linearly independent eigenvectors. (b) 3 linearly independent eigenvectors. (c) 4 linearly independent eigenvectors.

10.
(a) Symmetric matrix so diagonalizable.
(b) 2×2 Jordan block so not diagonalizable.
(c) 2 distinct eigenvalues so diagonalizable.
(d) $\lambda = 1$ with multiplicity 2 and only one eigenvector so not diagonalizable.

11. (a)

$$(i) \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{bmatrix}, \ (ii) \begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{bmatrix}, \ (iii) \begin{bmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}, \ (iv) \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & \\ & & & 2 \end{bmatrix}, \ (v) \begin{bmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{bmatrix}.$$

(b) (i): geometric multiplicity 4, (ii): geometric multiplicity 3, (iii): geometric multiplicity 2, (iv): geometric multiplicity 2, (v): geometric multiplicity 1.

(c)

$$(i) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (ii) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (iii) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad (iv) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad (v) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

(d) The minimal polynomials are (i) $q(\lambda) = \lambda - 2$, (ii) $q(\lambda) = (\lambda - 2)^2$, (iii) $q(\lambda) = (\lambda - 2)^2$, (iv) $q(\lambda) = (\lambda - 2)^3$, (v) $q(\lambda) = (\lambda - 2)^4$. Note that the Jordan forms in (ii) and (iii) have the same minimal polynomial.

12. $\lambda = 3$ is an eigenvalue of algebraic multiplicity 3 and $\text{rank}(A - 3I) = 1$ so the eigenvalue 3 has two eigenvectors associated with it: these are solution of $(A - 3I)x = 0$. We find that $x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $x_2 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$ are eigenvectors (and are linearly independent). To find the generalized eigenvector associated with $\lambda = 3$ solve $(A - 3I)v = x_1$ to get $v = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Now let $X = [\, x_1 \quad v \quad x_2 \,] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$. Check that $X^{-1}AX = \begin{bmatrix} 3 & 1 & \\ & 3 & \\ & & 3 \end{bmatrix}$.

13. The characteristic polynomial of A is $p(\lambda) = \lambda^2 - 3\lambda + 2$ and $p(A) = A^2 - 3A + 2I = O$. Thus $A^2 = 3A - 2I$, $A^3 = A(A^2) = 3A^2 - 2A = 3(3A - 2I) - 2A = 7A - 6I$, $A^4 = 7A^2 - 6A = 15A - 14I$.

14. Let $p(\lambda) = a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n$ be the characteristic polynomial of A with $a_0 \neq 0$ since A is nonsingular. From $p(A) = O$ we obtain $A^{-1}p(A) = a_0A^{-1} + a_1I + \cdots + a_{n-1}A^{n-2} + A^{n-1} = O$. Hence, $A^{-1} = (-1/a_0)(a_1I + \cdots + a_{n-1}A^{n-2} + A^{n-1})$.

15. By polynomial long division any such p can be written $p = q\,\psi + r$, where the degree of the remainder r is less than that of q. But $O = p(A) = q(A)\psi(A) + r(A) = r(A)$, and this contradicts the minimality of the degree of q unless $r = 0$. Hence $r = 0$ and q divides p.

16. Let $B = I - \frac{1}{3}A$. From $q(A) = O$ we obtain $5I + 9B^2 - 15B = O$. Premultiplying by B^{-1} and rearranging gives $B^{-1} = (3A + 6I)/5$, as required.

17. Note that $(uv^*)u = (v^*u)u$ so $\lambda = v^*u$ is an eigenvalue of uv^*. Let $v_2, \ldots, v_n$ be $n-1$ linearly independent vectors such that $v^*v_i = 0$, $i = 2, \ldots, n$. Then $(uv^*)v_i = 0$ so $\lambda = 0$ is an eigenvalue of multiplicity $n-1$. We deduce that the characteristic polynomial is $p(\lambda) = \lambda^{n-1}(\lambda - v^*u)$ and the minimal polynomial is $q(\lambda) = \lambda(\lambda - v^*u)$. As a check, we have $q(uv^*) = uv^*(uv^* - (v^*u)I) = v^*u(uv^* - uv^*) = O$.

18. It is easy to see that for scalars x and y, $p(x) - p(y) = q(x, y)(x - y)$ for some polynomial q of two variables. We can substitute tI for x and A for y to obtain

$$p(t)I - p(A) = q(tI, A)(tI - A). \tag{1}$$

If $p(A) = O$ then we have $p(t)(tI - A)^{-1} = q(tI, A)$, so that $p(t)(tI - A)^{-1}$ is a polynomial in t. Conversely, if $p(t)(tI - A)^{-1}$ is a polynomial in t then from (1) it follows that $p(A)(tI - A)^{-1} = p(t)(tI - A)^{-1} - q(tI, A)$ is a polynomial. Since $p(A)$ is a constant this implies that $p(A) = O$.

To obtain the Cayley–Hamilton theorem set $p(t) = \det(tI - A)$. From the formula $B^{-1} = \operatorname{adj}(B)/\det(B)$, where the adjugate adj is the transpose of the matrix of cofactors, we have $p(t)(tI - A)^{-1} = \operatorname{adj}(tI - A)$ is a polynomial in t, so $p(A) = O$ by the first part.

19.

(a) $A^kX = XB^k$ clearly holds for $k = 1$. Assume that $A^kX = XB^k$. Then $A^{k+1}X = A(A^kX) = AXB^k = XBB^k = XB^{k+1}$.

(b) From (a) it follows that $p(A)X = Xp(B)$ for any polynomial $p(t)$. Choose $p(t)$ to be the characteristic polynomial of A. By the Cayley–Hamilton theorem, $p(A) = O$ and therefore $p(A)X = O = Xp(B)$. Since $p(B) = (B - \lambda_1I)\cdots(B - \lambda_nI)$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A, the matrix $p(B)$ is nonsingular and $p(B)X = O$ has only the solution $X = O$.

Solution to **Example 5** from the handout:

$\underline{\lambda = 1}$: $\dim(\text{null}(A - I)) = n - \text{rank}(A - I) = 14 - 11 = 3$ so there are 3 Jordan blocks with eigenvalue $\lambda = 1$. Also since $\text{rank}(A - I)^3 = \text{rank}(A - I)^4$, the index of $\lambda = 1$ is 3 and therefore the size of the largest Jordan block with eigenvalue $\lambda = 1$ is 3. The formula in (4) (see handout) gives

number of blocks of size 1: $14 + 10 - 2 \times 11 = 2$.

number of blocks of size 2: $11 + 9 - 2 \times 10 = 0$.

number of blocks of size 3: $10 + 9 - 2 \times 9 = 1$.

Hence $\lambda = 1$ has algebraic multiplicity $2 \times 1 + 1 \times 3 = 5$.

$\underline{\lambda = 2}$: $\dim(\text{null}(A - 2I)) = n - \text{rank}(A - 2I) = 14 - 12 = 2$ so there are 2 Jordan blocks with eigenvalue $\lambda = 2$. Also since $\text{rank}(A - I)^2 = \text{rank}(A - I)^3$, the index of $\lambda = 2$ is 2 and therefore the size of the largest Jordan block with eigenvalue $\lambda = 2$ is 2. The formula in (4) gives

number of blocks of size 1: $14 + 10 - 2 \times 12 = 0$.

number of blocks of size 2: $12 + 10 - 2 \times 10 = 2$.

Hence $\lambda = 2$ has algebraic multiplicity $2 \times 2 = 4$.

$\underline{\lambda = 3}$: $\dim(\text{null}(A - I)) = n - \text{rank}(A - 3I) = 14 - 12 = 2$ so there are 3 Jordan blocks with eigenvalue $\lambda = 3$. Also since $\text{rank}(A - I)^4 = \text{rank}(A - I)^5$, the size of the largest Jordan block with eigenvalue $\lambda = 3$ is 4. The formula in (4) gives

number of blocks of size 1: $14 + 11 - 2 \times 12 = 1$.

number of blocks of size 2: $12 + 10 - 2 \times 11 = 0$.

number of blocks of size 3: $11 + 9 - 2 \times 10 = 0$.

number of blocks of size 4: $10 + 9 - 2 \times 9 = 1$.

Hence $\lambda = 1$ has algebraic multiplicity $1 \times 1 + 1 \times 4 = 5$.

From this information we obtain

$$J = \text{diag}\left(\begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix}, [1], [1], \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ & 2 \end{bmatrix}, \begin{bmatrix} 3 & 1 & & \\ & 3 & 1 & \\ & & 3 & 1 \\ & & & 3 \end{bmatrix}, [3] \right).$$

What is the minimal polynomial of this matrix?