Math 63CM Homework 5 Solutions

Kevin Yang

Stanford University

February 13, 2020

Problem 1

(i). Suppose $k = 0$; in this case, by the Cayley-Hamilton theorem we have coefficients $a_0, \ldots, a_{n-1}$ such that

$$A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 \mathrm{Id}_n = 0. \tag{0.1}$$

This proves the case for $k = 0$. For general k, multiplying the above equation by A^k gives a linear relation between $A^k, \ldots, A^{k+n}$, completing the proof.

(ii). Consider the matrix

$$A = (a_{ij})_{i,j=1}^n, \quad a_{ij} = \begin{cases} 1 & i = j-1 \\ 0 & i \neq j-1 \end{cases} \tag{0.2}$$

Namely, A is the matrix with 1s on the off-diagonal above the diagonal, and 0s everywhere else.

To prove that $\mathrm{Id}, A, \ldots, A^{n-1}$ are linearly independent, suppose we had a linear relation, i.e. coefficients $a_0, \ldots, a_{n-1}$ such that

$$a_0 \mathrm{Id} + a_1 A + \ldots + a_{n-1}A^{n-1} = 0. \tag{0.3}$$

Observe $A^n = 0$, but $A^{n-1} \neq 0$. Multiplying the previous relation by A^{n-1}, we thus obtain $a_0 A^{n-1} = 0$, so that $a_0 = 0$ necessarily.

Proceeding inductively, suppose we know $a_0, \ldots, a_j = 0$ for some $j = 0, \ldots, n-2$, so that

$$a_{j+1}A^{j+1} + \ldots + a_{n-1}A^{n-1} = 0. \tag{0.4}$$

Multiplying this relation by A^{n-2+j}, since $n \geqslant 2$, by the same token we obtain $a_{j+1}A^{n-1} = 0$, so that $a_{j+1} = 0$. This completes the proof.

Problem 2

(i). Suppose $(\lambda \mathrm{I} - A)^k v = 0$. Then clearly we have $(\lambda \mathrm{I} - A)^\ell v = 0$ if $\ell \geqslant k$, since we multiply the first relation by $(\lambda \mathrm{I} - A)^{\ell - k}$.

To show the reverse inclusion, if $(\lambda \mathrm{I} - A)^\ell v = 0$, we know $v \in V^{(\lambda)}$, where $V^{(\lambda)}$ is the generalized eigenspace of A with eigenvalue λ. However, we know that $(\lambda \mathrm{I} - A)$ is nilpotent as a map $V^{(\lambda)} \to V^{(\lambda)}$. In particular, if $k = \dim V^{(\lambda)}$, which is the assumption because k is the multiplicity of λ in the characteristic polynomial of A, we know that the characteristic polynomial of $(\lambda \mathrm{I} - A)$ is $p(\eta) = \eta^k$; this is because nilpotent maps only have eigenvalues of 0. By Cayley-Hamilton, we deduce $(\lambda \mathrm{I} - A)^k = 0$ as a map from $V^{(\lambda)} \to V^{(\lambda)}$. Thus, because $v \in V^{(\lambda)}$, we know $(\lambda \mathrm{I} - A)^k v = 0$ as well, which completes the proof.

(ii). Suppose $w \in \ker(A)$; we want to show $Bw \in \ker(A)$. To this end, by the commuting relation $AB = BA$, we have

$$ABw \;=\; BAw \;=\; B0 \;=\; 0 \tag{0.5}$$

which gives the claim. Switching A and B shows that if $v \in \ker(B)$, then $Av \in \ker(B)$.

It is not necessarily true that $\ker(A) = \ker(B)$. Consider $V = \mathbb{R}^2$ and

$$A \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B \;=\; \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{0.6}$$

We have $AB = BA = B$, but $\ker(B) = \mathbb{R}^2$ and $\ker(A) = 0$.

Problem 3

Suppose A is diagonalizable, first, so that we have a basis of eigenvectors $w_1, \ldots, w_n$ with eigenvalues $\lambda_{i_1}, \ldots, \lambda_{i_n} \subset \{\lambda_1, \ldots, \lambda_m\}$. Given any vector $v \in \mathbb{C}^n$, we write $v = \sum_{i=1}^n c_i w_i$. If we had $p(A)w_i = 0$ for all i, then of course $p(A)v = 0$, so it suffices to assume $v = w_i$ for some i. Then we have

$$p(A)w_i \;=\; \prod_{j=1}^{n} (A - \lambda_j \mathrm{I}) w_i. \tag{0.7}$$

However, we know the $(A - \lambda_j \mathrm{I})$-maps commute, so

$$p(A)w_i \;=\; \prod_{j \neq i} (A - \lambda_j \mathrm{I}) \cdot (A - \lambda_i \mathrm{I}) w_i \tag{0.8}$$

$$\;=\; 0, \tag{0.9}$$

since $Aw_i = \lambda_i w_i$.

Suppose now that $p(A)v = 0$ for all $v \in \mathbb{C}^n$. We use the Jordan canonical form for A, giving

$$\mathbb{C}^n \;\simeq\; V^{(\lambda_1)} \oplus \ldots \oplus V^{(\lambda_m)}, \tag{0.10}$$

where $V^{(\lambda_i)}$ is the λ_i-generalized eigenspace. If A is not diagonalizable, then for some i we may find $v \in V^{(\lambda_i)}$ such that $(A - \lambda_i)v \neq 0$.

Observe that because $\lambda_1, \ldots, \lambda_m$ are distinct, the map $(A - \lambda_j \mathrm{I})$ maps $V^{(\lambda_i)}$ to itself, and if $\lambda_j \neq \lambda_i$ this map is invertible on $V^{(\lambda_i)}$.

Now, by assumption and the commutativity mentioned before, we have

$$p(A)v \;=\; \prod_{j \neq i} (A - \lambda_j \mathrm{I}) \cdot (A - \lambda_i \mathrm{I}) v \;=\; 0. \tag{0.11}$$

However, by assumption $(A - \lambda_i \mathrm{I})v \neq 0$; moreover, by the invertibility of the $(A - \lambda_j \mathrm{I})$ on $V^{(\lambda_i)}$, we deduce $p(A)v \neq 0$; but this is a contradiction.

Problem 4

We claim

$$\det(\lambda \mathrm{I} - \mathrm{e}^A) \;=\; \prod_{i=1}^{m} (\lambda - \mathrm{e}^{\lambda_i})^{\nu_i}. \tag{0.12}$$

Consider the Jordan canonical form for A, so that $\mathbb{C}^n \simeq V^{(\lambda_1)} \oplus \ldots \oplus V^{(\lambda_m)}$. We now claim that e^{λ_i} is an eigenvalue of e^A, and that $V^{(\lambda_i)} \subseteq V^{(\mathrm{e}^{\lambda_i})}$, where the latter is the generalized eigenspace for e^A with eigenvalue e^{λ_i}, for each i. To show this, it suffices to show that $\mathrm{e}^A - \mathrm{e}^{\lambda_i}\mathrm{I}$ is nilpotent as a map from $V^{(\lambda_i)}$ to itself; first, observe this map actually sends $V^{(\lambda_i)}$ to itself, because e^A is a sum of powers of A, and A maps $V^{(\lambda_i)}$ to itself.

We now compute

$$
\begin{aligned}
\mathrm{e}^A - \mathrm{e}^{\lambda_i}\mathrm{I} &= \sum_{k=1}^{\infty} \frac{A^k - \lambda_i^k \mathrm{I}}{k!} && (0.13)\\
&= (A-\lambda_i \mathrm{I}) \sum_{k=1}^{\infty} \frac{A^{k-1} + \lambda_i A^{k-2} + \ldots + \lambda_i^{k-2} A + \lambda_i^{k-1}\mathrm{I}}{k!}. && (0.14)
\end{aligned}
$$

Because the sum commutes with the factor $(A-\lambda_i\mathrm{I})$, because $(A-\lambda_i\mathrm{I})$ is nilpotent as a map on $V^{(\lambda_i)}$, we deduce $\mathrm{e}^A - \mathrm{e}^{\lambda_i}\mathrm{I}$ is as well.

We now have

$$
\mathbb{C}^n \simeq V^{(\lambda_i)} \oplus \ldots \oplus V^{(\lambda_m)} \subseteq V^{(\mathrm{e}^{\lambda_1})} \oplus \ldots \oplus V^{(\mathrm{e}^{\lambda_m})} \subseteq \mathbb{C}^n, \tag{0.15}
$$

the last inclusion being a consequence of the map given by the Jordan canonical form for e^A. Counting dimensions, by rank-nullity we deduce all inclusions are isomorphisms. In particular, the Jordan canonical form of e^A is given by $V^{(\mathrm{e}^{\lambda_1})} \oplus \ldots \oplus V^{(\mathrm{e}^{\lambda_m})} \simeq \mathbb{C}^n$, where $\dim V^{(\mathrm{e}^{\lambda_i})} = \dim V^{\lambda_i} = \nu_i$ for all i. This provides the desired characteristic polynomial for e^A.

Problem 5

(i). The characteristic polynomial of this matrix is $p(\lambda) = (\lambda^2+1)^2 = (\lambda - i)^2(\lambda+i)^2$. Thus, to find the diagonalizable operator L, it suffices to find a basis for the generalized eigenspaces $V^{(i)}$ and $V^{(-i)}$. This will actually resolve part (ii) if we choose the right bases.

First, we observe that $V^{(i)}$ and $V^{(-i)}$ are both 2-dimensional.

- We first find a good basis for the $\lambda = i$ generalized eigenspace. To this end, we must solve the following equation for $v \in \mathbb{C}^4$:

$$
\begin{pmatrix} 1-i & -1 & 0 & 1 \\ 2 & -1-i & 1 & 0 \\ 0 & 0 & -1-i & 2 \\ 0 & 0 & -1 & 1-i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{0.16}
$$

Solving this system of equations, we obtain $v = \begin{pmatrix} 1+i \\ 2 \\ 0 \\ 0 \end{pmatrix}$, or any scalar multiple of it.

To complete the basis for $V^{(i)}$, let us find a vector w such that $Aw = iw + v$. Thus, $\{v, w\}$ would be a basis for the $V^{(i)}$-generalized eigenspace. In particular, we need to solve

$$
\begin{pmatrix} 1-i & -1 & 0 & 1 \\ 2 & -1-i & 1 & 0 \\ 0 & 0 & -1-i & 2 \\ 0 & 0 & -1 & 1-i \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} 1+i \\ 2 \\ 0 \\ 0 \end{pmatrix}. \tag{0.17}
$$

One can check that $w = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1+i \end{pmatrix}$ is a solution. This gives our basis $\{v, w\}$.

- To find a good basis for the $\lambda = -i$-generalized eigenspace, you can do the same procedure. You can also realize taking complex conjugates of everything in the first bullet point turns into the problem we have to solve now, so we get a basis $\{u, z\}$ with $u = \begin{pmatrix} 1-i \\ 2 \\ 0 \\ 0 \end{pmatrix}$ and $z = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1-i \end{pmatrix}$.

Thus, the desired diagonalizable map L is given by

$$L = S \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} S^{-1}, \quad S = \begin{pmatrix} 1+i & 0 & 1-i & 0 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 1+i & 0 & 1-i \end{pmatrix}. \tag{0.18}$$

Of course, one can multiply this out and get $L = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & -1 & 1 \end{pmatrix}$, which is certainly a matrix with real-valued entries! If you didn't get a real-entried matrix, something went wrong.

(ii). Here, the desired matrix is just S from above.

PROBLEM 6

(i). It suffices to show that for some $c > 0$ and all $x \in \mathbb{R}^n$ with $\|x\| = 1$, we have $x \cdot Ax > c$. Indeed, if this were true, for any arbitrary nonzero $x \in \mathbb{R}^n$, for $z = \|x\|^{-1}x$ we would have $z \cdot Az > c$, from which we deduce that $x \cdot Ax > c\|x\|^2$.

To prove the claim for unit vectors, observe the map $x \mapsto x \cdot Ax$ is continuous (by Cauchy-Schwarz, for example). Thus, it achieves a minimum on the compact unit sphere, say at x_0. But for x with $\|x\| = 1$, we then have $x \cdot Ax \geqslant \inf_{x:\|x\|=1} x \cdot Ax = x_0 \cdot Ax_0 > 0$, so that we may take $c = x_0 \cdot Ax_0$.

(ii). For any nonzero eigenvector v with eigenvalue λ, we have $v \cdot Av = \lambda\|v\|^2$. However, by part (i), we know $\lambda\|v\|^2 > c\|v\|^2$, so that $\lambda > c > 0$.

(iii). Because A is real symmetric, for some orthogonal matrix O we have $A = ODO^T$, where D is diagonal with diagonal entries $\lambda_1, \dots, \lambda_n$ the positive eigenvalues of A. We then have

$$\int_0^T e^{-tA}\, dt = \int_0^T Oe^{-tD}O^T\, dt \tag{0.19}$$

$$= O\int_0^T e^{-tD}\, dt O^T. \tag{0.20}$$

However, e^{-tD} is diagonal with i-th diagonal entry $e^{-t\lambda_i}$. In particular, integrating entrywise, the i-th entry of $\int_0^T e^{-tD}\, dt$ is given by

$$\int_0^T e^{-t\lambda_i}\, dt = \lambda_i^{-1}\left[1 - e^{-T\lambda_i}\right] \tag{0.21}$$

since $\lambda_i > 0$. Taking the limit as $T \to \infty$, the RHS converges to λ_i^{-1}, so that

$$\lim_{T\to\infty} \int_0^T e^{-tA}\, dt = OD^{-1}O^T. \tag{0.22}$$

This actually proves the result desired in part (iv) as well!

(iv). See above.

PROBLEM 7

(i). By definition, we have

$$\begin{aligned}\langle Av, w\rangle &= \sum_{i=1}^{n}[Av]_i\bar{w}_i &(0.23)\\ &= \sum_{i,j=1}^{n} a_{ij}v_j\bar{w}_i &(0.24)\\ &= \sum_{j=1}^{n} v_j\overline{\sum_{i=1}^{n}\bar{a}_{ij}w_i} &(0.25)\\ &= \sum_{j=1}^{n} v_j\overline{\sum_{i=1}^{n}a_{ji}w_i} &(0.26)\\ &= \sum_{j=1}^{n} v_j\overline{[Aw]_j} &(0.27)\\ &= \langle v, Aw\rangle. &(0.28)\end{aligned}$$

(ii). Suppose we have an eigenvalue λ of A; consider a nonzero eigenvector v with this eigenvalue. We have

$$\langle Av, v\rangle = \sum_{i=1}^{n}\lambda v_i\bar{v}_i = \lambda\|v\|^2. \quad (0.29)$$

Observe $\|v\|^2$ is real and nonzero. Moreover, by part (i), we have

$$\begin{aligned}\langle Av, v\rangle &= \langle v, Av\rangle &(0.30)\\ &= \sum_{i=1}^{n} v_i\cdot\bar{\lambda}\bar{v}_i &(0.31)\\ &= \bar{\lambda}\sum_{i=1}^{n} v_i\bar{v}_i &(0.32)\\ &= \bar{\lambda}\|v\|^2. &(0.33)\end{aligned}$$

Thus, we have $\lambda\|v\|^2 = \bar{\lambda}\|v\|^2$, so that $\lambda = \bar{\lambda}$ and thus λ is real.

(iii). Differentiating, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n} v_i(t)\overline{v_i(t)} &(0.34)\\ &= \sum_{i=1}^{n} v_i'(t)\cdot\overline{v_i(t)} + \sum_{i=1}^{n} v_i(t)\cdot\overline{v_i'(t)} &(0.35)\\ &= \langle\frac{\mathrm{d}}{\mathrm{d}t}v(t),\ v(t)\rangle + \langle v(t),\ \frac{\mathrm{d}}{\mathrm{d}t}v(t)\rangle &(0.36)\\ &= \langle i^{-1}Av(t), v(t)\rangle + \langle v(t),\ i^{-1}Av(t)\rangle. &(0.37)\end{aligned}$$

However, because $i^{-1} = -i$, we have $\overline{i^{-1}} = i = -i^{-1}$. Thus, we have $\langle v(t),\ i^{-1}Av(t)\rangle. = \langle -i^{-1}v(t), Av(t)\rangle$, and because A is self-adjoint, we have

$$\langle i^{-1}Av(t), v(t)\rangle + \langle v(t),\ i^{-1}Av(t)\rangle = 0. \quad (0.38)$$

Thus, we know $\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|^2 = 0$, which completes the proof.

(iv). We do the same thing; using the calculation in part (iii), this gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|^2 \;=\; \langle \frac{\mathrm{d}}{\mathrm{d}t}v(t),\; v(t)\rangle \;+\; \langle v(t),\; \frac{\mathrm{d}}{\mathrm{d}t}v(t)\rangle \tag{0.39}$$

$$\;=\; \langle i^{-1}D(t)v(t), v(t)\rangle \;+\; \langle -i^{1}v(t), D(t)v(t)\rangle. \tag{0.40}$$

Indeed, the A-terms cancel as in part (iii). However, we know that $D(t)$ is a symmetric matrix with real-valued entries. In particular, we know

$$\langle -i^{1}v(t), D(t)v(t)\rangle \;=\; \langle -i^{1}D(t)v(t), v(t)\rangle \tag{0.41}$$

and thus

$$\langle i^{-1}D(t)v(t), v(t)\rangle \;+\; \langle -i^{1}v(t), D(t)v(t)\rangle \;=\; \langle i^{-1}D(t)v(t), v(t)\rangle \;+\; \langle -i^{1}D(t)v(t), v(t)\rangle \tag{0.42}$$

$$\;=\; 0. \tag{0.43}$$

This again completes the proof.