Matrix Analysis Homework 7

Michael Nelson

Problem a

Problem a.1

Since V has dimension n, the set $\{v, f(v), \ldots, f^{n-1}(v), f^n(v)\}$ (which as size $n+1$) is linearly dependent. Therefore there exists $c_0, c_1, \ldots, c_{n-1}, c_n \in K$ (not all equal to 0) such that

$$c_0 v + c_1 f(v) + \cdots + c_{n-1} f^{n-1}(v) + c_n f^n(v) = 0. \tag{1}$$

Choose such $c_0, c_1, \ldots, c_{n-1}, c_n \in K$.

Assume (for a contradiction) that $c_n = 0$. Then we can rewrite (1) as

$$c_0 v + c_1 f(v) + \cdots + c_{n-1} f^{n-1}(v) = 0. \tag{2}$$

Since $\{v, f(v), \ldots, f^{n-1}(v)\}$ is linearly independent, (2) implies $c_0 = c_1 = \cdots = c_{n-1} = 0$. But this is a contradiction since $c_0, c_1, \ldots, c_{n-1}, c_n$ are not all equal to 0. Thus $c_n \neq 0$, and hence we can rewrite (1) as

$$f^n(v) = a_{n-1} f^{n-1}(v) + \cdots + a_1 f(v) + a_0 v.$$

where $a_i = -c_i/c_n$ for each $0 \leq i \leq n-1$.

Problem a.2

If $\dim V = 1$, then the matrix representation of f with respect to $\{v\}$ is just the 1×1 matrix (a_0), so assume $\dim V > 1$. Let $\beta := \{v, f(v), \ldots, f^{n-1}(v)\}$. Since

$$f(f^i(v)) = f^{i+1}(v)$$

for all $0 \leq i < n-1$ and

$$f(f^{n-1}(v)) = a_{n-1} f^{n-1}(v) + \cdots + a_1 f(v) + a_0 v,$$

we have

$$[f]_\beta = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{n-1} \end{pmatrix}$$

Problem a.3

We first solve the following problem: for each $n \in \mathbb{N}$, let $c_0, c_1, \ldots, c_n \in K$ and let

$$P_{c_0, c_1, \ldots, c_n}(X) := \det \begin{pmatrix} -X & 0 & \cdots & 0 & c_0 \\ 1 & -X & \cdots & 0 & c_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & -X & c_{n-1} \\ 0 & 0 & \cdots & 1 & c_n - X \end{pmatrix}.$$

We prove by induction on $n \geq 1$ that

$$P_{c_0, c_1, \ldots, c_n}(X) = (-1)^{n+1}(X^{n+1} - c_n X^n - \cdots - c_1 X - c_0). \tag{3}$$

For the base case $n = 1$, we have

$$\begin{aligned}
P_{c_0,c_1}(X) &= \det \begin{pmatrix} -X & c_0 \\ 1 & c_1 - X \end{pmatrix} \\
&= -X(c_1 - X) - c_0 \\
&= X^2 - c_1 X - c_0 \\
&= (-1)^2 (X^2 - c_1 X - c_0)
\end{aligned}$$

Now let $n > 1$ and assume that (3) is true for all $k < n$. Then for any $c_0, c_1, \dots, c_n \in K$, we have

$$\begin{aligned}
P_{c_0,c_1,\dots,c_n}(X) &= \det \begin{pmatrix} -X & 0 & \cdots & 0 & c_0 \\ 1 & -X & \cdots & 0 & c_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & -X & c_{n-1} \\ 0 & 0 & \cdots & 1 & c_n - X \end{pmatrix} \\
&= -X \det \begin{pmatrix} -X & \cdots & 0 & c_1 \\ 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & -X & c_{n-1} \\ 0 & \cdots & 1 & c_n - X \end{pmatrix} + (-1)^n c_0 \det \begin{pmatrix} 1 & -X & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -X \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\
&= -X \det \begin{pmatrix} -X & \cdots & 0 & c_1 \\ 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & -X & c_{n-1} \\ 0 & \cdots & 1 & c_n - X \end{pmatrix} + (-1)^n c_0 \\
&= -X P_{c_1,\dots,c_n}(X) + (-1)^n c_0 \\
&= -X((-1)^n (X^n - c_n X^{n-1} - c_{n-1} X^{n-2} - \cdots - c_1)) + (-1)^n c_0 \\
&= (-1)^{n+1} (X^{n+1} - c_n X^n - c_{n-1} X^{n-1} - \cdots - c_1 X - c_0).
\end{aligned}$$

where we used the induction step to get from the fourth line to the fifth line.

Now we find the characteristic polynomial of f. If $\dim V = 1$, then the characteristic polynomial of f is given by $\chi_f(X) = a_0 - X$, so assume $\dim V > 1$. Then characteristic polynomial of f is given by

$$\begin{aligned}
\chi_f(X) &= \det([f]_\beta - X\mathrm{I}_n) \\
&= \det \begin{pmatrix} -X & 0 & \cdots & 0 & a_0 \\ 1 & -X & \cdots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & -X & \vdots \\ 0 & 0 & \cdots & 1 & a_{n-1} - X \end{pmatrix} \\
&= P_{a_0,a_1,\dots,a_{n-1}}(X) \\
&= (-1)^n (X^n - a_{n-1} X^{n-1} - \cdots - a_1 X - a_0).
\end{aligned}$$

Problem b

We give V the structure of a $K[X]$-module by defining

$$p(X) \cdot v = p(f)(v) \tag{4}$$

for all $p(X) \in K[X]$ and for all $v \in V$. That (4) does indeed give V the structure of a $K[X]$-module is shown in the Appendix.

Problem b.1

Let $v, w \in \ker(p(X))$ and let $a, b \in K$. Then

$$\begin{aligned}
p(X) \cdot (av + bw) &= p(f)(av + bw) \\
&= \sum_{i=0}^{n} c_i f^i(av + bw) \\
&= \sum_{i=0}^{n} c_i(af^i(v) + bf^i(w)) \\
&= a \sum_{i=0}^{n} c_i f^i(v) + b \sum_{i=0}^{n} c_i f^i(w) \\
&= a(p(X) \cdot v) + b(p(X) \cdot w) \\
&= 0 + 0 \\
&= 0.
\end{aligned}$$

Thus $av + bw \in \ker(p(X))$ which implies $\ker(p(X))$ is a linear subspace of V. In particular, when $p(X) = X - \lambda$ where $\lambda \in K$, we have

$$\begin{aligned}
v \in \ker(p(X)) &\iff v \in \ker(X - \lambda) \\
&\iff (X - \lambda) \cdot v = 0 \\
&\iff (f - \lambda)(v) = 0 \\
&\iff f(v) = \lambda v.
\end{aligned}$$

Thus $v \in \ker(p(X))$ if and only if v is an eigenvector of f with eigenvalue λ. Therefore $\ker(p(X)) = E_\lambda$ where E_λ is the eigenspace of f with respect to λ.

Problem b.2

Write

$$p(X) = \sum_{i=0}^{m} c_i X^i \quad \text{and} \quad q(X) = \sum_{j=0}^{n} d_j X^j$$

We first show that

$$\ker(p(X)q(X)) = \ker(p(X)) + \ker(q(X)). \tag{5}$$

Let $v \in \ker(p(X)) + \ker(q(X))$. Write $v = v_1 + v_2$ where $v_1 \in \ker(p(X))$ and $v_2 \in \ker(q(X))$. Then

$$\begin{aligned}
(p(X)q(X)) \cdot v &= p(X) \cdot (q(X) \cdot v) \\
&= p(X) \cdot (q(X) \cdot (v_1 + v_2)) \\
&= p(X) \cdot (q(X) \cdot v_1 + q(X) \cdot v_2) \\
&= p(X) \cdot (q(X) \cdot v_1) \\
&= (p(X)q(X)) \cdot v_1 \\
&= (q(X)p(X)) \cdot v_1 \\
&= q(X) \cdot (p(X) \cdot v_1) \\
&= q(X) \cdot 0 \\
&= 0
\end{aligned}$$

implies $v \in \ker(p(X)q(X))$. Thus $\ker(p(X)) + \ker(q(X)) \subseteq \ker(p(X)q(X))$. For the reverse inclusion, choose $a(X), b(X) \in K[X]$ so that

$$a(X)p(X) + b(X)q(X) = 1. \tag{6}$$

Let $v \in \ker(p(X)q(X))$. Using (6), write $v = v_1 + v_2$ where

$$v_1 = (b(X)q(X)) \cdot v \quad \text{and} \quad v_2 = (a(X)p(X)) \cdot v.$$

Then $v_2 \in \ker(q(X))$ since

$$\begin{aligned}
q(X) \cdot v_2 &= q(X) \cdot ((a(X)p(X)) \cdot v) \\
&= (q(X)a(X)p(X)) \cdot v \\
&= (a(X)p(X)q(X)) \cdot v \\
&= a(X) \cdot (p(X)q(X) \cdot v) \\
&= a(X) \cdot 0 \\
&= 0.
\end{aligned}$$

Similarly, $v_1 \in \ker(p(X))$ since

$$\begin{aligned} p(X) \cdot v_1 &= p(X) \cdot ((b(X)q(X)) \cdot v) \\ &= (p(X)b(X)q(X)) \cdot v \\ &= (b(X)p(X)q(X)) \cdot v \\ &= b(X) \cdot (p(X)q(X) \cdot v) \\ &= b(X) \cdot 0 \\ &= 0. \end{aligned}$$

Therefore $v \in \ker(p(X)) + \ker(q(X))$, and this implies $\ker(p(X)) + \ker(q(X)) \supseteq \ker(p(X)q(X))$.
To see that (5) is a direct sum, let $v \in \ker(p(X)) \cap \ker(q(X))$. Then

$$\begin{aligned} v &= 1 \cdot v \\ &= (a(X)p(X) + b(X)q(X)) \cdot v \\ &= (a(X)p(X)) \cdot v + (b(X)q(X)) \cdot v \\ &= a(X) \cdot (p(X) \cdot v) + b(X) \cdot (q(X) \cdot v) \\ &= a(X) \cdot 0 + b(X) \cdot 0 \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Thus $\ker(p(X)) \cap \ker(q(X)) = 0$ and so the sum (5) is direct.

Problem b.3

We first prove by induction on $m \geq 2$ that for polynomials $p_i(X) \in K[X]$ such that $\gcd(p_i(X), p_j(X)) = 1$ for all $1 \leq i < j \leq m$, we have

$$\ker(p_1(X)p_2(X) \cdots p_m(X)) = \ker(p_1(X)) \oplus \ker(p_2(X)) \oplus \cdots \oplus \ker(p_m(X)). \tag{7}$$

The base case $m = 2$ was established in problem b.2. Now assume (7) is true for some $m \geq 2$. Let $p_i(X) \in K[X]$ such that $\gcd(p_i(X), p_j(X)) = 1$ for all $1 \leq i < j \leq m+1$. Since $\gcd(p_1(X), p_i(X)) = 1$ for all $2 \leq i \leq m+1$, we have $\gcd(p_1(X), p_2(X) \cdots p_{m+1}(X)) = 1$. Therefore

$$\begin{aligned} \ker(p_1(X)p_2(X) \cdots p_{m+1}(X)) &= \ker(p_1(X)) \oplus \ker(p_2(X) \cdots p_{m+1}(X)) \\ &= \ker(p_1(X)) \oplus \ker(p_2(X)) \oplus \cdots \oplus \ker(p_{m+1}(X)), \end{aligned}$$

where we used the base case on the first line and where we used the induction hypothesis to get from the first line to the second line.
To finish the problem, we just need to show that $V = \ker(c(X))$. Let $v \in V$. Then

$$\begin{aligned} c(X) \cdot v &= c(f)(v) \\ &= 0(v) \\ &= 0 \end{aligned}$$

implies $v \in \ker(c(X))$. Therefore $V \subseteq \ker(c(X))$, which implies $V = \ker(c(X))$ (since $\ker(c(X))$ was already shown to be a subspace of V in problem b.1).

Problem b.4

Let $E = \sum_{i=1}^{t} E_{\lambda_i}$ and let $c(X)$ be given by

$$c(X) = (X - \lambda_1) \cdots (X - \lambda_t),$$

where $\lambda_1, \ldots, \lambda_t$ are the distinct eigenvalues of f. Since $(X - \lambda_i)$ and $(X - \lambda_j)$ are relatively prime for all $1 \leq i < j \leq t$ and since $c(f) = 0$ on E, we can apply problem b.3 and obtain

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_t}$$

In particular $B_1 \cup B_2 \cup \cdots \cup B_t$ must be linearly independent: Suppose

$$\sum_{i=1}^{t} \sum_{j=1}^{m_i} a_{ij} u_{ij} = 0. \tag{8}$$

Then for each $1 \leq i \leq t$, we must have $\sum_{j=1}^{m_i} a_{ij}u_{ij} = 0$. Indeed, if $\sum_{j=1}^{m_k} a_{kj}u_{kj} \neq 0$ for some $1 \leq k \leq t$, then we can rearrange (8) to get

$$\sum_{j=1}^{m_k} a_{kj}u_{kj} = -\sum_{\substack{1\leq i\leq t \\ i\neq k}} \sum_{j=1}^{m_i} a_{ij}u_{ij},$$

and so

$$\begin{aligned} 0 \neq \sum_{j=1}^{m_k} a_{kj}u_{kj} \\ &\in E_{\lambda_k} \cap \bigoplus_{\substack{1\leq i\leq t \\ i\neq k}} E_{\lambda_i} \\ &= \{0\}, \end{aligned}$$

gives us our desired contradiction. Thus, for each $1 \leq i \leq t$, we have

$$\sum_{j=1}^{m_i} a_{ij}u_{ij} = 0.$$

But this implies $a_{ij} = 0$ for all $1 \leq j \leq m_i$ since B_i is a basis for all $1 \leq i \leq t$. Thus $a_{ij} = 0$ for all $1 \leq i \leq t$ and $1 \leq j \leq m_i$, and hence $B_1 \cup B_2 \cup \cdots \cup B_t$ is linearly independent.

Appendix

$K[X]$-module

Let us check that the action (4) does indeed give V the structure of a $K[X]$-module. Obviously V is an abelian group since it is a K-vector space. Also we have $1 \cdot v = v$ for all $v \in V$, where 1 is the identity in $K[X]$. Let $p(X), q(X) \in K[X]$ and let $v, w \in V$. Write

$$p(X) = \sum_{i=0}^{m} c_i X^i \quad \text{and} \quad q(X) = \sum_{j=0}^{n} d_j X^j.$$

Then

$$\begin{aligned} (p(X) + q(X)) \cdot v &= (p(f) + q(f))(v) \\ &= \left(\sum_{i=0}^{m} c_i f^i + \sum_{j=0}^{n} d_j f^j \right)(v) \\ &= \sum_{i=0}^{m} c_i f^i(v) + \sum_{j=0}^{n} d_j f^j(v) \\ &= p(f)(v) + q(f)(v) \\ &= p(X) \cdot v + q(X) \cdot v \end{aligned}$$

and

$$\begin{aligned} p(X) \cdot (v + w) &= p(f)(v + w) \\ &= \sum_{i=0}^{m} c_i f^i(v + w) \\ &= \sum_{i=0}^{m} c_i (f^i(v) + f^i(w)) \\ &= \sum_{i=0}^{m} c_i f^i(v) + \sum_{i=0}^{m} c_i f^i(w) \\ &= p(f)(v) + p(f)(w) \\ &= p(X) \cdot v + p(X) \cdot w \end{aligned}$$

and

$$\begin{aligned}
p(X) \cdot (q(X) \cdot v) &= p(X) \cdot (q(f)(v)) \\
&= p(X) \cdot \sum_{j=0}^{n} d_j f^j(v) \\
&= \sum_{j=0}^{n} d_j (p(X) \cdot f^j(v)) \\
&= \sum_{j=0}^{n} d_j p(f)(f^j(v)) \\
&= \sum_{j=0}^{n} d_j \left(\sum_{i=0}^{m} c_i f^i (f^j(v)) \right) \\
&= \sum_{j=0}^{n} d_j \sum_{i=0}^{m} c_i f^{i+j}(v) \\
&= \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} c_i d_{k-i} \right) f^k(v) \\
&= (p(X)q(X)) \cdot v.
\end{aligned}$$

Thus all of the required properties for V to be a $K[X]$-module under the action (4) are satisfied.