Prof. M. Einsiedler
Prof. P. Biran

Musterlösung Serie 18

Diagonalizabiltiy, Cayley-Hamilton

1. Let K be a field, let $n \geqslant 2$, and let
$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ 0 & 0 & \cdots & 0 & -c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{n-1} \end{pmatrix} \in M_{n\times n}(K).$$
Prove that
$$\mathrm{char}_A(X) = (-1)^n (X^n + c_{n-1}X^{n-1} + \cdots + c_0).$$
Hint: Use induction.

Solution: First note that for $n = 2$:
$$\det \begin{pmatrix} -X & -c_0 \\ 1 & -X - c_1 \end{pmatrix} = X^2 + c_1 X + c_0 = (-1)^2 (X^2 + c_1 X + c_0).$$
Now assume that $n \geqslant 3$ and that we have shown the statement for $n-1$. Expand the determinant with repsect to the first row to obtain
$$\begin{aligned} &\det \begin{pmatrix} -X & 0 & \cdots & 0 & -c_0 \\ 1 & -X & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ 0 & 0 & \cdots & 0 & -c_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -X - c_{n-1} \end{pmatrix} \\ = \quad & (-X) \det \begin{pmatrix} -X & 0 & \cdots & 0 & -c_1 \\ 1 & -X & \cdots & 0 & -c_2 \\ 0 & 1 & \cdots & 0 & -c_3 \\ 0 & 0 & \cdots & 0 & -c_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -X - c_{n-1} \end{pmatrix} + (-1)^{n+1}(-c_0) \\ = \quad & (-X)(-1)^{n-1}(X^{n-1} + c_{n-1}X^{n-2} + \cdots + c_1) + (-1)^n c_0 \\ = \quad & (-1)^n (X^n + c_{n-1}X^{n-1} + \cdots + c_0), \end{aligned}$$
which concludes the proof.

2. Let A be a $n \times n$-matrix of rank r. Show that the degree of the minimal polynomial of A is smaller or equal than $r+1$.

   Solution: By the definition of rank, the image of L_A is a subspace of dimension r. Consider the restriction

   $$F := L_A|_{\mathrm{Bild}(L_A)} : \mathrm{Bild}(L_A) \to \mathrm{Bild}(L_A).$$

   The characteristic polynomial $q(X) := \mathrm{char}_F(X)$ of F has degree r, and by Cayley Hamilton, we have $q(F) = 0$.

   For $p(X) := q(X) \cdot X$, we get $v \in K^n$

   $$p(L_A)(v) = (q(L_A) \circ L_A)(v) = q(L_A)(Av) = q(F)(Av) = 0,$$

   and hence $p(L_A) = 0$. Thus, we have $p(A) = 0$. By definition, the minimal polynomial of A divides $p(X)$. Therefore the degree of the first is $\leqslant$ the degree of $p(X)$, hence less or equal $r+1$.

3. Prove that every 2×2 invertible real matrix belongs to one of the following categories:

   - It is diagonalizable;
   - it is trigonalizable with algebraic multiplicity 2 and geometric multiplicity 1;
   - one can find a basis such that the matrix representation in that basis is

   $$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \quad \text{with } b \neq 0.$$

   Solution: First note that the characteristic polynomial of any such matrix, denoted p, is a quadratic polynomial with real coefficients. If all of its roots are real, it splits into linear factors in $\mathbb{R}[X]$. In this case, we either have

   - a diagonalizable matrix if the algebraic multiplicity equals the algebraic multiplicity for each root (see the first theorem of the lecture notes eigenvectors.c, in English) or
   - a trigonalizable matrix (see the first theorem on page 4 of the same lecture notes).

   Let us now address the case where the polynomial does not split into linear factors over $\mathbb{R}$. Then, by the fundamental theorem of algebra, it must admit a complex root λ. Moreover, if we write $p(X) = X^2 + c_1 X + c_0$, we have that $p(\lambda) = 0$ implies

   $$0 = \overline{p(\lambda)} = \overline{\lambda^2 + c_1\lambda + c_0} = (\overline{\lambda})^2 + c_1\overline{\lambda} + c_0 = p(\overline{\lambda}),$$

   since the coefficients of p are real. Hence the complex conjugate of λ is the other root of p. We are left to show that every such characteristic polynomial arises from a rotation matrix. We prove the following result:

Theorem. *Let A be a 2×2 real matrix with a complex eigenvalue $\lambda \in \mathbb{C} \setminus \mathbb{R}$, and let v be a complex eigenvector corresponding to λ. Then, $A = CBC^{-1}$ for*

$$C = \begin{pmatrix} | & | \\ \Re(v) & \Im(v) \\ | & | \end{pmatrix} \quad \textit{and} \quad B = \begin{pmatrix} \Re(\lambda) & \Im(\lambda) \\ -\Im(\lambda) & \Re(\lambda) \end{pmatrix},$$

where

$$\Re \begin{pmatrix} x+iy \\ z+iw \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} \quad \textit{and} \quad \Im \begin{pmatrix} x+iy \\ z+iw \end{pmatrix} = \begin{pmatrix} y \\ w \end{pmatrix}.$$

In other words, C is similar to a rotation-scaling matrix.

Beweis. We first need to show that the vectors $\Re(v)$ and $\Im(v)$ are linearly independent in order to prove that C is invertible. If not, there exist $x, y \in \mathbb{R}$ such that $x\Re(v) + y\Im(v) = 0$. Then

$$\begin{aligned}(y+ix)v &= y\Re(v) - x\Im(v) + i(x\Re(v) + y\Im(v)) \\ &= y\Re(v) - x\Im(v) \in \mathbb{R}^2.\end{aligned}$$

Hence, on one hand $y\Re(v) - x\Im(v)$ is a complex multiple of v, therefore it is an eigenvector of A with eigenvalue λ. On the other hand, it is a real eigenvector of A, and since A is real, it can only correspond to a real eigenvalue. This is a contradiction.

Let us denote $\lambda = a + ib$ and $v = \left(\begin{smallmatrix} x+iy \\ z+iw \end{smallmatrix}\right)$. Since $\{\Re(v), \Im(v)\}$ forms a basis of $\mathbb{R}^2$,

$$CBC^{-1} = A \iff \begin{cases} A\Re(v) &= CBC^{-1}\Re(v) \\ A\Im(v) &= CBC^{-1}\Im(v) \end{cases}$$

On one hand,

$$\begin{aligned} A\Re(v) + iA\Im(v) &= Av \\ &= \lambda v \\ &= (a+ib)\begin{pmatrix} x+iy \\ z+iw \end{pmatrix} \\ &= \begin{pmatrix} ax-by \\ az-bw \end{pmatrix} + i\begin{pmatrix} bx+ay \\ bz+aw \end{pmatrix}. \end{aligned}$$

On the other hand,

$$Ce_1 = \Re(v) \Leftrightarrow e_1 = C^{-1}\Re(v) \Leftrightarrow CBe_1 = CBC^{-1}\Re(v)$$

and similarly replacing e_1 by e_2 and $\Re(v)$ by $\Im(v)$. Hence

$$CBC^{-1}\Re(v) = CBe_1 = C\begin{pmatrix} a \\ -b \end{pmatrix} = \begin{pmatrix} ax-by \\ az-bw \end{pmatrix} = A\Re(v)$$

and similarly $CBC^{-1}\Im(v) = A\Im(v)$. □

4. Let K be a field, $A \in M_{n\times n}(K)$, and $p \in K[X]$ be a non-trivial polynomial such that $p(A) = 0$. Show that every eigenvalue of A is a root of p.

   Hint: For an eigenvector v of A and a polynomial q over K, state and prove the relationship between v and $q(A) \cdot v$.

   Solution: We proceed as indicated in the hint and consider $q(A)\cdot v$ for a non-trivial polynomial
$$q(X) = b_n X^n + b_{n-1}X^{n-1} + \cdots + b_0 \in K[X]$$
   and an eigenvector v of A. We have

$$\begin{aligned} q(A) \cdot v &= b_n A^n \cdot v + b_{n-1}A^{n-1} \cdot v + \cdots + b_0 I_n \cdot v \\ &= b_n \lambda^n v + b_{n-1}\lambda^{n-1} v + \cdots + b_0 v \\ &= (b_n \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_0)v \end{aligned}$$

   Now assume that $p(X) = c_n X^n + c_{n-1}X^{n-1} + \cdots + c_0 \in K[X] \smallsetminus \{0\}$ is such that $p(A) = 0$. We have

$$(c_n \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0)v = p(A) \cdot v = 0.$$

   Hence $c_n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0 = 0$, i.e. λ is a root of p.

5. (a) Let A be a $n \times n$-matrix. Prove that the subspace $\langle I_n, A, A^2, \ldots \rangle$ of $M_{n\times n}(K)$ has dimension $\leqslant n$.

   (b) Let $A := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$. Find a polynomial $p(X)$ with $p(A) = A^{-1}$.

   Solution:

   (a) *Solution*: The subspace $W := \langle I_n, A, A^2, \ldots, A^{n-1} \rangle$ of $M_{n\times n}(K)$ is generated by n elements and thus has dimension $\leqslant n$. Thus, it is enough to show that $W = \langle I_n, A, A^2, \ldots \rangle$.The inclusion „$\subset$“ is already clear, hence we need to show

   Claim: For all $k \geqslant 0$, we have $A^k \in W$.

   We prove by induction on k.

   Induction start: For $k \leqslant n - 1$ this holds by construction of W.

   Oops: In the case of $n = 0$ this claim is empty, hence no start. But then $M_{n\times n}(K)$ is the zero space and the space in question as well, hence has dimension $n = 0$, as desired. Now let $n \geqslant 1$.

   Induction step: Let $k \geqslant n$.

   Induction hypothesis: The Claim is true for all smaller values of k.

The characteristic polynomial of A is monic of degree n; write it as $\mathrm{char}_A(X) = X^n + \sum_{i=0}^{n-1} a_i X^i$. By Cayley-Hamilton, we have $\mathrm{char}_A(A) = 0$, thus

$$A^n = -\sum_{i=0}^{n-1} a_i A^i.$$

Multiplying with A^{k-n} yields

$$A^k = -\sum_{i=0}^{n-1} a_i A^{i+k-n} = -\sum_{j=k-n}^{k-1} a_{j-k+n} A^j.$$

By induction hypothesis all multiples of A^j on the right side are contained in W; thus A^k also lies in W, which we wanted to show.

(b) The characteristic polynomial of A is

$$\mathrm{char}_A(X) = -X^3 + 6X^2 + 3X - 18$$

By Cayley-Hamilton, we have

$$\mathrm{char}_A(A) = -A^3 + 6A^2 + 3A - 18I_3 = 0,$$

also

$$I_3 = -\frac{1}{18}A^3 + \frac{1}{3}A^2 + \frac{1}{6}A = \left(-\frac{1}{18}A^2 + \frac{1}{3}A^2 + \frac{1}{6}I_3\right) \cdot A.$$

In particular, we get that A is invertible and

$$A^{-1} = -\frac{1}{18}A^2 + \frac{1}{3}A + \frac{1}{6}I_3.$$

6. Prove or disprove: There exists a real $n \times n$-matrix A satisfying

$$A^2 + 2A + 5I_n = 0$$

if and only if n is even.

Solution: The claim is true!

„$\Rightarrow$“: The real polynoial $p(X) := X^2 + 2X + 5$ has the complex zeros $-1 \pm 2i$; these are not real. Now let A be a $n \times n$-matrix with $p(A) = 0$ and n odd. Then the characteristic polynomial of A had odd degree n, and thus has a real zero λ. This is an eigenvalue of A, with corresponding Eigenvector v. For this we have

$$0 = p(A)v = (A^2 + 2A + 5I_n)v = (\lambda^2 + 2\lambda + 5)v = P(\lambda)v,$$

and hence $P(\lambda) = 0$. This is a contradiction to p not having a real zero.

„$\Leftarrow$“: By Cayley-Hamilton, every 2×2-matrix A_2 with characteristic polynomial $p(X)$ satisfies the desired equation, for example the companion matrix $A_2 := \left(\begin{smallmatrix} -2 & 1 \\ -5 & 0 \end{smallmatrix}\right)$. For arbitrary $n \geqslant 0$ let A be the $n \times n$-matrix

$$A := \begin{pmatrix} A_2 & & \\ & \ddots & \\ & & A_2 \end{pmatrix},$$

then we have

$$A^2 + 2A + 5I_n = \begin{pmatrix} A_2 + 2A_2 + 5I_2 & & \\ & \ddots & \\ & & A_2 + 2A_2 + 5I_2 \end{pmatrix} = 0.$$