Math 217 Worksheet 24: The Characteristic Polynomial (§7.2, §7.3)

Fix a linear transformation $V \xrightarrow{T} V$, where V is an n-dimensional vector space.
Definition. The **characteristic polynomial** of T is

$$\chi_T(x) = \det(T - xI_V)$$

where $\det(T - xI_V)$ denotes the determinant of the linear transformation $V \xrightarrow{S} V$ such that $S(v) = T(v) - xv$ for any $v \in V$. Here, we can think of x as an indeterminate real number.

Theorem A. The characteristic polynomial is a polynomial of degree n in x (where $n = \dim V$).

Theorem B. The (real) eigenvalues of T are the (real) *roots* of the characteristic polynomial.

Corollary. The transformation T has at most n distinct eigenvalues (where $n = \dim V$).

Problem 1. Let $V \xrightarrow{T} V$ be a linear transformation, where V is a finite dimensional vector space.

(a) Discuss with your group how to *define* the determinant of T. How can you get a *matrix* out of T? Why is the determinant of this matrix independent of any choices you made?

(b) Now let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be reflection over the line spanned by $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

(i) Compute $[T]_\mathcal{E}$, $[xI_V]_\mathcal{E}$, and $[T - xI_V]_\mathcal{E}$ (where x is an indeterminate scalar).

(ii) Thinking geometrically, find an eigenbasis $\mathcal{B}$ for T. Compute $[T]_\mathcal{B}$, $[xI_V]_\mathcal{B}$, $[T - xI_V]_\mathcal{B}$.

(iii) Compute and compare the determinants of $[T - xI_V]_\mathcal{E}$ and $[T - xI_V]_\mathcal{B}$. What is the determinant of $[T - xI_V]_\mathcal{C}$ where $\mathcal{C}$ is some third basis?

(iv) Compute the polynomial $\chi_T(x)$ and confirm that its roots are the eigenvalues of T.

Solution:

(a) To define the determinant of T, we pick *any* basis $\mathcal{B}$. This allows us to model V as the $\mathcal{B}$-coordinate space $\mathbb{R}^n$ (where $n = \dim V$) and T by the matrix multiplication by the $n \times n$ matrix $[T]_\mathcal{B}$. The determinant of $[T]_\mathcal{B}$ is the determinant of T. It doesn't depend of the choice of basis $\mathcal{B}$, since two matrices representing T in different basis are *similar*, meaning that $[T]_\mathcal{B} = S[T]_\mathcal{A}S^{-1}$ where S is a change of basis matrix. By the multiplicative property of determinants, we see $\det[T]_\mathcal{B} = \det S \det[T]_\mathcal{A} \det S^{-1} = \det[T]_\mathcal{A}$. Since the characteristic polynomial is a determinant, it doesn't depend on choice of basis used to compute it.

(b) (i) $[T]_\mathcal{E} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$, $[xI_V]_\mathcal{E} = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}$, and $[T - xI_V]_\mathcal{E} = \begin{bmatrix} -x & -1 \\ -1 & -x \end{bmatrix}$.

(ii) An eigenbasis $\mathcal{B}$ for T is $\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}.\}$ In this basis, we have $[T]_\mathcal{B} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $[xI_V]_\mathcal{B} = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}$, and $[T - xI_V]_\mathcal{B} = \begin{bmatrix} 1-x & 0 \\ 0 & -1-x \end{bmatrix}$.

(iii) All are $x^2 - 1$.

(iv) $\chi_T(x) = x^2 - 1$. Its roots are ± 1, the eigenvalues of T.

Problem 2. Find the matrix $[T - xI_V]_{\mathcal{B}}$ and the characteristic polynomial of each transformation below, for some *conveniently chosen basis* $\mathcal{B}$. Then use Theorem B to find all eigenvalues of each.

(a) $T : \mathcal{P}_2 \to \mathcal{P}_2$ is differentiation.*

(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ is projection onto the plane with normal vector $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

(c) $T : \mathbb{R}^2 \to \mathbb{R}^2$ whose standard matrix is $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$.

(d) $T : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ defined by $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & a+b \\ a+b+c & a+b+c+d \end{bmatrix}$.

Solution:

(a) Using the basis $\mathcal{B} = (1, x, x^2)$, the $\mathcal{B}$-matrix of $T - xI_V$ is $\begin{bmatrix} -x & 1 & 0 \\ 0 & -x & 2 \\ 0 & 0 & -x \end{bmatrix}$ and the characteristic polynomial is $-x^3$. So 0 is the only eigenvalue.

(b) Using a basis $\mathcal{B}$ consisting of the normal vector $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ together with a spanning set for the plane, the $\mathcal{B}$-matrix of $T - xI_V$ is $\begin{bmatrix} -x & 0 & 0 \\ 0 & 1-x & 0 \\ 0 & 0 & 1-x \end{bmatrix}$ and the characteristic polynomial is $-x(x-1)^2$. The eigenvalues are the roots—so 0 and 1.

(c) Using the standard basis, we have $[T - xI_V]_{\mathcal{E}}$ is $A - xI_2 = \begin{bmatrix} 1-x & 2 \\ -1 & 4-x \end{bmatrix}$. Its char poly is $(1-x)(4-x) + 2 = x^2 - 5x + 6$. This factors as $(x-2)(x-3)$ so the eigenvalues are 2 and 3.

(d) $T : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ defined by $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & a+b \\ a+b+c & a+b+c+d \end{bmatrix}$. Using the basis $\mathcal{E} = (E_{11}, E_{12}, E_{21}, E_{22})$, the matrix $T - xI$ is $\begin{bmatrix} 1-x & 0 & 0 & 0 \\ 1 & 1-x & 0 & 0 \\ 1 & 1 & 1-x & 0 \\ 1 & 1 & 1 & 1-x \end{bmatrix}$. The char poly is $(1-x)^4 = (x-1)^4$. The only eigenvalue is 1.

Definition. Let $f(x)$ be a polynomial, and let λ be a root. The **multiplicity of the root** λ is the largest m so that we can factor $f(x) = (x-\lambda)^m g(x)$ for some $m \in \mathbb{N}$ for some $g(x)$.

Definition. Let λ be an eigenvalue of a linear transformation T of a finite dimensional space. The **algebraic multiplicity** of λ is its multiplicity as a root of the characteristic poly $\chi_T(x)$.

*Save your answers to (a) and all other parts of Problem 2! You will need them for Problem 4.

Problem 3.

(a) Find the algebraic multiplicity (or "almu") of each eigenvalue in each part of Problem 2.

(b) Prove that the Corollary follows from Theorems A and B.
[Hint: You may assume basic facts about roots of polynomials of degree n learned in high school or Math 115.]

(c) Prove that for *any* linear transformation $T : V \to V$, with V a finite dimensional vector space, the sum of the algebraic multiplicities of the eigenvalues of T is at most $\dim V$.

Solution:

(a) We factor each characteristic polynomial into linear factors to see the multiplicities of each root. For (a): $\chi(x) = -x^3$, so $\text{almu}(0) = 3$.
For (b): $\chi(x) = -x(x-1)^2$, so $\text{almu}(1) = 2$. $\text{almu}(0) = 1$.
For (c): $\chi(x) = (x-2)(x-3)$, $\text{almu}(2) = 1$. $\text{almu}(3) = 1$.
For (d): $\chi(x) = (x-1)^4$, so $\text{almu}(1) = 4$.

(b) Remembering that a polynomial of degree n has at most n real roots, it follows immediately that a linear transformation $T : V \to V$ of an n dimensional space V has at most n eigenvalues. Indeed, the eigenvalues are precisely the roots of the Char poly (by Theorem B), which has degree n (by Theorem A).

(c) Let $\chi(x)$ be the characteristic polynomial for a linear transformation $T : V \to V$ of an n dimensional space V. Factor $\chi(x)$ as much as possible:

$$\chi(x) = (x-\lambda_1)^{a_1}(x-\lambda_2)^{a_2}\cdots(x-\lambda_d)^{a_d}g(x)$$

where $\lambda_1, \ldots, \lambda_d$ are *all* the eigenvalues of T (of algebraic multiplicities $a_1, \ldots, a_d$, respectively) and $g(x)$ is a polynomial with *no linear factors*. Note that the degree of the polynomial χ is $n = a_1 + \ldots a_d + \deg(g)$. So $\sum_{i=1}^{d} \text{almu}(\lambda_i) \leq n$.

Theorem C. Let $T : V \to V$ be a linear transformation of a finite dimensional vector space V. For each eigenvalue λ of T, the geometric multiplicity of λ is *at most* the algebraic multiplicity of λ. For short, we write "$\text{gemu}(\lambda) \leq \text{almu}(\lambda)$."

Problem 4. For each transformation in Problem 2, find the eigenspace of each eigenvalue. Then confirm, in these examples, that Theorem C holds.

Solution: To find the geometric multiplicity of λ, we need the dimension of the corresponding eigenspace. The eigenspace is modeled by the kernel of the matrix $[T - xI_V]_{\mathcal{B}}$ with the specific value λ plugged in for x.

(a) For T as in part (a) of Problem 2, the kernel of differentiation is just the constant functions, a space of dimension one. So $E_0 = \mathcal{P}_0$, and $\text{gemu}(0) = 1$. As we saw in Problem 3, $\text{almu}(0) = 3$, so we've conformed $\text{gemu}(0) \leq \text{almu}(0)$ for this map.. Alternatively, we can compute the geometric multiplicity using rank-nullity: the 0-eigenspace is modeled by the kernel of the matrix $[T - 1I_V]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$. By rank-nullity, we know the kernel, and hence the 0-eigenspace is one dimensional, and $\text{gemu}(0) = 1$.

(b) To find the eigenspaces for the map in Problem 2b, we can argue geometrically: the vectors sent to zero by the projection are normal to the plane, so $E_0 = \text{Span}(\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix})$; the vectors sent to themselves by the projection are on the plane, so $E_1 =$ the plane. Note that here gemu$(1) = 2$. gemu$(0) = 1$. The almu and gemu of each eigenvalue are equal for this T.

You can also find the eigenspaces algebraically, by finding the kernels of $\begin{bmatrix} -x & 0 & 0 \\ 0 & 1-x & 0 \\ 0 & 0 & 1-x \end{bmatrix}$ after plugging in $x = 0$ or $x = 1$.

(c) The 2-eigenspace is the kernel of $A - 2I_2 = \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix}$, which is the span of $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$. So gemu(2)=almu(2)=1.

The 3-eigenspace is the kernel of $A - 3I_2 = \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix}$, which is the span of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. So gemu(3)=almu(3)=1.

(d) The 1-eigenspace is represented by the kernel of the matrix $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$. Since this matrix is rank three, its kernel is 1-dimensional. We can take $\vec{e}_4$ as a basis, so E_{22} spans the 1-eigenspace of T. The geometric multiplicity is one, while the algebraic multiplicity is 4.

Problem 5: Proof of Theorem B. Fix $A \in \mathbb{R}^{n\times n}$ and $\lambda \in \mathbb{R}$. Show the following are equivalent:

(i) λ is an eigenvalue of the transformation $T_A : \mathbb{R}^n \to \mathbb{R}^n$ given by $T_A(\vec{v}) = A(\vec{v})$.

(ii) The matrix $A - \lambda I_n$ has nullity greater than zero.

(iii) The matrix $A - \lambda I_n$ has rank less than n.

(iv) The matrix $A - \lambda I_n$ has determinant zero.

(v) λ is a root of the characteristic polynomial $\det(A - xI_n)$.

Now explain why Theorem B follows.

Solution: We've mostly already done this on WS 23. The scalar λ is an eigenvalue for T if and only if the kernel of $A - \lambda I_n$ is not zero (since this kernel is the λ-eigenspace when it is non-zero). That is, (i) and (ii) are equivalent. Note (ii) and (iii) are equivalent by rank-nullity. Because the matrix $A - \lambda I_n$ is square, its kernel is non-zero if and only if its determinant is zero, so (iv) is also equivalent. Of course, (v) is just the definition of the characteristic polynomial. Theorem B follows: modeling T in any basis $\mathcal{B}$, we know λ is an eigenvalue of T if and only if λ is an eigenvalue of $[T]_{\mathcal{B}}$ and T and $[T]_{\mathcal{B}}$ have the same characteristic polynomial, by definition.

Theorem D. Eigenvectors of *distinct* eigenvalues are linearly independent. That is, if $T : V \to V$ is a linear transformation, and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ are eigenvectors with *different* eigenvalues, then $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a linearly independent subset of V.

Problem 6. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation in Problem 2b. Let $\vec{v}$ and $\vec{w}$ be eigenvectors of T with *different* eigenvalues. Verify Theorem D for the set $\{\vec{v}, \vec{w}\}$. [HINT: Think geometrically!]

Solution: In Problem 2a, we have only 0 and 1 as the distinct eigenvalues. The zero eigenvectors lie in Span($\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$), while the 1-eigenvectors lie in $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}^{\perp}$. So every zero-eigenvector is orthogonal to every 1-eigenvector. Since orthogonal vectors are linearly independent, we conclude that if $\vec{v}$ and $\vec{w}$ are eigenvectors of T with different eigenvalues, then $\{\vec{v}, \vec{w}\}$ are linearly independent.

Problem 7. Let $T : V \to V$ be a linear transformation. Using Theorem D, prove that:

(a) If T has n *distinct eigenvalues* where $n = \dim V$, then T has an eigenbasis. Find an example where this occurs from among the transformations in Problem 2.

(b) If $A \in \mathbb{R}^{n\times n}$ has n *distinct eigenvalues*, then A is similar to a diagonal matrix. For the matrix A in Problem 2c, find a diagonal matrix D and an invertible matrix S such that $A = SDS^{-1}$.

(c) The converses of (a) and (b) are *false*. Find two counterexamples to (b) when $n = 3$. One example should have only one eigenvalue and the other two. Use your examples to disprove (a) as well.

Solution:

(a) If there are n-distinct eigenvalues, the corresponding n eigenvectors are linearly independent by Theorem D. Since these are n linearly independent vectors in an n dimensional space, they form a basis (Theorem A WS12). Thus, they form an eigenbasis.

(b) If the matrix A has n distinct eigenvalues, then the corresponding transformation has an eigenbasis $\mathcal{B}$. In this basis, the matrix is diagonal and all other matrices for the transformation, including A (which is the matrix in the standard basis) are similar to it. The matrix S is the change of basis matrix. For example, with A as in Problem 2c, we have $A = SDS^{-1}$ where $S = S_{\mathcal{B}\to\mathcal{E}} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is diagonalizable and has only the eigenvalue 1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ has two eigenvalues, 0 and 1.

Problem 8. Let $\lambda_1, \lambda_2, \ldots, \lambda_d$ be *distinct* eigenvalues of a linear transformation $T : V \to V$.

(a) Say that v_1 is an λ_1-eigenvector and v_2 is an λ_2-eigenvector. *Without using Theorem D*, prove that v_1 is not a scalar multiple of v_2.

(b) Prove Theorem D. [HINT: Induce on d. Assume a relation $c_1v_1 + \cdots + c_dv_d = 0$. Now apply T and suitably subtract to find a relation on fewer vectors.]

Solution:

(a) If v_2 is a scalar multiple of v_1, then since $v_1 \in E_{\lambda_1}$, and E_{λ_1} is a vector space, also v_2 is in E_1. But then v_2 has eigenvalue λ_1, not λ_2.

(b) We induce on d. If $d=1$, the statement is trivial since eigenvectors are never zero, so $\{v_1\}$ is a linearly independent set. Part (a) takes care of the case where $d=2$. Assume, inductively, that any set of $d-1$ eigenvectors with distinct eigenvalues are linearly independent. Suppose, by way of contradiction, that there is a non-trivial relation on the set $\{v_1,\ldots,v_d\}$. Without loss of generality, we can write $a_1v_1+\cdots+a_nv_n=0$ with all a_i non-zero (by re-ordering the v_i's if needed). If $n<d$, we are done by induction. So assume $n=d$. Apply T. We have $0=T(a_1v_1+\cdots+a_nv_n)=T(a_1v_1)+\cdots+T(a_nv_n)=a_1\lambda_1v_1+\cdots+a_n\lambda_n\vec{v}_n$. Here the λ are the eigenvalues, and by assumption, no two are equal. Note that no $\lambda_i=0$, or else we already have a non-trivial relation with fewer non-zero terms. Now multiply the original relation by λ_1 and subtract.

We get $0=(\lambda_1a_2-\lambda_2a_2)v_2+\cdots+(\lambda_1a_n-\lambda_na_n)v_n=0$. This is a relation on $n-1$ (which is $d-1$) eigenvectors with distinct eigenvalues, so it must be trivial, by induction. So each $\lambda_1a_i-\lambda_ia_i=0$. But this also gives a contradiction, since $a_i\neq 0$ and $\lambda_1\neq\lambda_i$. QED

Problem 9*. Prove Theorem A. [Hint: If you have done the Extra WS on "Patterns and Determinants," the easiest way uses the pattern definition of determinant. Otherwise, use induction and the multilinearity of the determinant.]

Solution: By fixing a basis, it suffices to show that for $A\in\mathbb{R}^{n\times n}$, $\det(A-xI_n)$ is a polynomial of degree n in x. For a proof using the "pattern" definition of determinant, see the book; proof of Theorem 7.2.5.

We give a different proof, by induction on n.

Base Case: When $n=1$, $A=[a]$ and $\det(A-xI_1)=\det[a-x]=(a-x)$. This is a degree 1 polynomial in x.

Inductive Step: Assume that for $B\in\mathbb{R}^{n-1\times n-1}$, $\det(B-xI_{n-1})$ is a polynomial of degree $n-1$ in x. Now, for $A\in\mathbb{R}^{n\times n}$, $\det(A-xI_n)$, we consider the determinant as a linear function in column one to get

$$\det(A-xI_n)=\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22}-x & a_{23} & \cdots & a_{2n}\\ \vdots & & & \ddots & \vdots\\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}-x\end{bmatrix}+\det\begin{bmatrix} -x & a_{12} & a_{13} & \cdots & a_{1n}\\ 0 & a_{22}-x & a_{23} & \cdots & a_{2n}\\ \vdots & & & \ddots & \vdots\\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn}-x\end{bmatrix}.$$

Consider only the second summand for a moment: computing its determinant using a Laplace expansion down the first column, we have

$$\det\begin{bmatrix} -x & a_{12} & a_{13} & \cdots & a_{1n}\\ 0 & a_{22}-x & a_{23} & \cdots & a_{2n}\\ \vdots & & & \ddots & \vdots\\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn}-x\end{bmatrix}=-x\det\begin{bmatrix} a_{22}-x & a_{23} & \cdots & a_{2n}\\ \vdots & & \ddots & \vdots\\ a_{n2} & a_{n3} & \cdots & a_{nn}-x\end{bmatrix}=-x\chi_B,$$

where B is the $(n-1)\times(n-1)$ matrix obtained by deleting the first row and column of A. By induction, we can assume that χ_B is a polynomial of degree $n-1$, so this summand is a polynomial of degree n. To complete the proof, it suffices to show that the other summand,

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22}-x & a_{23} & \cdots & a_{2n}\\ \vdots & & & \ddots & \vdots\\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}-x\end{bmatrix},$$

is a polynomial of degree at most $n-1$. This can be accomplished by induction as well. When $n-1=1$, we have

$$\det\begin{bmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}-x\end{bmatrix}=a_{11}(a_{22}-x)-a_{12}a_{21},$$

which is a polynomial of degree zero or one in x, depending on whether $a_{11} = 0$ or not. Then computing the determinant by Laplace expansion down the first column, we have

$$\begin{aligned} &\det \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & a_{23} & \cdots & a_{2n} \\ \vdots & & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} - x \end{bmatrix} \\ &\quad = a_{11}(\det B - xI_{n-1}) - a_{21} \det B_{21} - a_{31} \det B_{31} + \cdots + (-1)^{n+1} a_{n1} \det B_{n1} \end{aligned} \tag{1}$$

where B is the same $(n-1) \times (n-1)$ matrix as above and each B_{i1} is an $(n-1) \times (n-1)$ matrix of the form

$$\begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1n} \\ b_{22} & b_{23} - x & \cdots & b_{2n} \\ \vdots & & \ddots & \vdots \\ b_{n2} & b_{n2} & \cdots & b_{nn} - x \end{bmatrix}.$$

By induction, each of the summands in (1) is a polynomial of degree at most $n-1$. QED.