MATH1030 Characteristic polynomial of a matrix.

1. **Definition. (Characteristic polynomial of a matrix.)**

   Let A be an $(n \times n)$-square matrix. The (algebraic) expression $\det(A - xI_n)$ *(with indeterminate x) is called the characteristic polynomial of the matrix A, and is denoted by $p_A(x)$.*

2. **Examples.**

   (a) Suppose $A = \begin{bmatrix} 13 & 30 \\ -6 & -14 \end{bmatrix}$. Then

   $$\begin{aligned} p_A(x) &= \det(A - xI_2) = \det(\begin{bmatrix} 13-x & 30 \\ -6 & -14-x \end{bmatrix}) \\ &= (13-x)(-14-x) - (-6)\cdot 30 = x^2 + x - 2. \end{aligned}$$

   Observations:

   - $p_A(x)$ is a degree-2 polynomial with leading coefficient 1 and constant term $\det(A)$.
   - $p_A(x)$ can be factorized as $p_A(x) = (x-1)(x+2)$. Coincidentally, the real roots of $p_A(x)$ are the eigenvalues of A.

     We have $A\mathbf{u} = 1 \cdot \mathbf{u}$, and $A\mathbf{v} = -2\mathbf{v}$, where $\mathbf{u} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

   (b) Suppose $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$. Then

   $$\begin{aligned} p_A(x) &= \det(A - xI_3) = \det(\begin{bmatrix} 1-x & 1 & 1 \\ 0 & 2-x & 2 \\ 0 & 0 & 3-x \end{bmatrix}) \\ &= (1-x)(2-x)(3-x) = -(x-1)(x-2)(x-3) = -x^3 + 6x^2 - 11x + 6. \end{aligned}$$

   Observations:

   - $p_A(x)$ is a degree-3 polynomial with leading coefficient -1 and constant term $\det(A)$.
   - $p_A(x)$ can be factorized as $p_A(x) = -(x-1)(x-2)(x-3)$. Coincidentally, the real roots of $p_A(x)$ are the eigenvalues of A.

     We have $A\mathbf{u} = 1 \cdot \mathbf{u}$, $A\mathbf{v} = 2\mathbf{v}$ and $A\mathbf{w} = 3\mathbf{w}$, where $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$.

   (c) Suppose $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$. Then

   $$\begin{aligned} p_A(x) &= \det(A - xI_3) = \det(\begin{bmatrix} 2-x & 1 & 1 \\ 1 & 2-x & 1 \\ 1 & 1 & 2-x \end{bmatrix}) = \det(\begin{bmatrix} 2-x & 1 & 1 \\ 1 & 2-x & 1 \\ 0 & -1+x & 1-x \end{bmatrix}) \\ &= \det(\begin{bmatrix} 2-x & 2 & 1 \\ 1 & 3-x & 1 \\ 0 & 0 & 1-x \end{bmatrix}) \\ &= (1-x)\det(\begin{bmatrix} 2-x & 2 \\ 1 & 3-x \end{bmatrix}) = (1-x)\det(\begin{bmatrix} 2-x & 2 \\ -1+x & 1-x \end{bmatrix}) = (1-x)\det(\begin{bmatrix} 4-x & 2 \\ 0 & 1-x \end{bmatrix}) \\ &= (1-x)^2(4-x) = -(x-1)^2(x-4) = -x^3 + 6x^2 - 9x + 4. \end{aligned}$$

   Observations:

   - $p_A(x)$ is a degree-3 polynomial with leading coefficient -1 and constant term $\det(A)$.
   - $p_A(x)$ can be factorized as $p_A(x) = -(x-1)^2(x-4)$. Coincidentally, the real roots of $p_A(x)$ are the eigenvalues of A.

     We have $A\mathbf{u} = 4\mathbf{u}$, $A\mathbf{v}_1 = 1 \cdot \mathbf{v}$ and $A\mathbf{v}_2 = 1 \cdot \mathbf{v}_2$, where $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

   (d) Suppose $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & 1 & 1 & 1 \\ -5 & 2 & 5 & -1 \end{bmatrix}$.

   Then $p_A(x) = \det(A - xI_4) = \cdots = (x+3)(x+1)(x-1)(x-3)$. (Fill in the calculations.)

   Observations:

- $p_A(x)$ is a degree-4 polynomial with leading coefficient 1 and constant term $\det(A)$.
- $p_A(x)$ can be factorized as $p_A(x) = (x+3)(x+1)(x-1)(x-3)$. Coincidentally, the real roots of $p_A(x)$ are the eigenvalues of A.

  We have $A\mathbf{t} = 1 \cdot \mathbf{t}$, $A\mathbf{u} = -1 \cdot \mathbf{u}$, $A\mathbf{v} = 3 \cdot \mathbf{v}$, $A\mathbf{w} = -3 \cdot \mathbf{w}$, where $\mathbf{t} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 5 \\ -1 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 1 \\ -5 \\ -3 \\ 15 \end{bmatrix}$.

(e) Let b be a real number. Suppose $A = \begin{bmatrix} b & 1 & 0 \\ 0 & b & 1 \\ 0 & 0 & b \end{bmatrix}$. Then

$$p_A(x) = \det(A - xI_3) = \det(\begin{bmatrix} b-x & 1 & 0 \\ 0 & b-x & 1 \\ 0 & 0 & b-x \end{bmatrix}) = (b-x)^3 = -x^3 + 3bx^2 - 3b^2x + b^3$$

Observations:

- $p_A(x)$ is a degree-3 polynomial with leading coefficient -1 and constant term $\det(A)$.
- $p_A(x)$ can be factorized as $p_A(x) = -(x-b)^3$.
  The only (real) root of $p_A(x)$ is the only eigenvalue of A.

  We have $A\mathbf{u} = b\mathbf{u}$, where $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

(f) Suppose $A = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$. Then

$$\begin{aligned} p_A(x) &= \det(A - xI_4) = \det(\begin{bmatrix} 1-x & 0 & 0 & -1 \\ 1 & 1-x & 0 & 0 \\ 0 & 1 & 1-x & 0 \\ 0 & 0 & 1 & 1-x \end{bmatrix}) \\ &= (1-x)\det(\begin{bmatrix} 1-x & 0 & 0 \\ 1 & 1-x & 0 \\ 0 & 1 & 1-x \end{bmatrix}) - \det(\begin{bmatrix} 0 & 0 & -1 \\ 1 & 1-x & 0 \\ 0 & 1 & 1-x \end{bmatrix}) \\ &= (1-x)^4 + (1-x)\det(\begin{bmatrix} 0 & -1 \\ 1 & 1-x \end{bmatrix}) = (1-x)^4 + 1 = x^4 - 4x^3 + 6x^2 - 4x + 2 \end{aligned}$$

Observations:

- $p_A(x)$ is a degree-4 polynomial with leading coefficient -1 and constant term $\det(A)$.
- $p_A(x)$ has no real roots. Coincidentally, A has no eigenvalues.

3. **Theorem (1).**

   Suppose A is an $(n \times n)$-square matrix.

   Then $p_A(x)$ is a degree-n polynomial with indeterminate x, with leading coefficient $(-1)^n$, and with constant coefficient $\det(A)$.

   Remark. The multiple of $(-1)^{n-1}$ with the coefficient of the degree-$(n-1)$ term in the polynomial $p_A(x)$ is called the trace of A, and is denoted by $\mathrm{tr}(A)$.

   Proof of Theorem (1). Omitted. (This is an exercise in mathematical induction.)

4. Recall that a square matrix is singular if and only if its determinant is zero. As a consequence of this logical equivalence, we have the result below:

   Theorem (E).

   Suppose A is an $(n \times n)$-square matrix, and λ is a real number. Then the statements below are logically equivalent:

   (a) *λ is an eigenvalue of A.*

   (b) *$A - \lambda I_n$ is singular.*

   (c) $\det(A - \lambda I_n) = 0$.

   (d) *λ is a real root of $p_A(x)$.*

Remark. Now suppose λ is indeed an eigenvalue of A. So λ is a real root of $p_A(x)$ indeed. According to the Factor Theorem, $p_A(x) = (x-\lambda)f(x)$ for some polynomial with real coefficients $f(x)$. Repeatedly applying the Factor Theorem, we can show that there is some uniquely determined positive integer m_λ for which $p_A(x) = (x-\lambda)^m_\lambda g(x)$ for some polynomial with real coefficients $g(x)$ and for which $p_A(x)$ is not divisible by $(x-\lambda)^{m_\lambda+1}$. Such an integer m is called the algebraic multiplicity of the eigenvalue λ of A. It can be shown that $\dim(\mathcal{E}_A(\lambda)) \leq m_\lambda$.

5. Note that every polynomial of odd degree and with real coefficients has at least one real root. Then we have the result below:

   Theorem (2).

   Let A be an $(n \times n)$-square matrix. Suppose n is odd. Then A has at least one eigenvalue.

6. **Theorem (3).**

   Suppose A is a symmetric (2×2)-square matrix. Then A is diagonalizable.

   Proof of Theorem (3).

   Suppose A is a symmetric (2×2)-square matrix. Then $A = \begin{bmatrix} a_1 & c \\ c & a_2 \end{bmatrix}$ for some real numbers a_1, a_2, c.

   Write $\alpha = \dfrac{a_1 + a_2}{2}$, and $\beta = \dfrac{a_1 - a_2}{2}$. Note that $\alpha^2 - \beta^2 = a_1 a_2$.

   We have $p_A(x) = \det(A - xI_2) = (a_1 - x)(a_2 - x) - c^2 = x^2 - (a_1 + a_2)x + a_1 a_2 - c^2 = x^2 - 2\alpha x + \alpha^2 - \beta^2 - c^2 = (x-\alpha)^2 - (\beta^2 + c^2) = (x - \alpha - \sqrt{\beta^2 + c^2})(x - \alpha + \sqrt{\beta^2 + c^2})$.

   Then $p_A(x)$ has two (not necessarily) distinct real roots, namely $\alpha + \sqrt{\beta^2 + c^2}$, $\alpha - \sqrt{\beta^2 + c^2}$.

   - (Case 1.) Suppose the two real roots of $p_A(x)$ are distinct. Then A is diagonalizable by Theorem (C).
   - (Case 2.) Suppose the two real roots of $p_A(x)$ are the same number. Then $\beta^2 + c^2 = 0$.

     Therefore $\beta = c = 0$. Hence $A = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$.

     So A is a diagonal matrix. It is trivially diagonalizable.

   Hence, in any case, A is diagonalizable.

7. Theorem (3) is a special case of Theorem (F), whose proof is beyond the scope of this course. (The easiest argument is given through complex numbers.)

   Theorem (F).

   Suppose A is a symmetric $(n \times n)$-square matrix. Then A is diagonalizable.

   Illustrations.

   (a) Let $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

   Note that A is symmetric. Then we expect A to be diagonalizable by Theorem (E).

   In fact, a diagonalization for A given by $U^{-1}AU = \mathrm{diag}(4, 1, 1)$, with $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$, and $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

   (b) Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

   Note that A is symmetric. Then we expect A to be diagonalizable by Theorem (E).

   In fact, a diagonalization for A given by $U^{-1}AU = \mathrm{diag}(2, -1, -1)$, with $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \mathbf{u}_3\ ]$, and $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$.

8. **Theorem (4).**

   Suppose A is a diagonalizable $(n \times n)$-square matrix.

For each j, denote the coefficient of the j-th power term of $p_A(x)$ is c_j. (So $p_A(x) = c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1}+c_nx^n$ as polynomials.)

Then $c_0I_n + c_1A + c_2A^2 + \cdots + c_{n-1}A^{n-1} + c_nA^n = \mathcal{O}_{n\times n}$.

Remark. The conclusion in Theorem (4) is often presented as $p_A(A) = \mathcal{O}_{n\times n}$.

9. Theorem (4) is a special case of the result below, whose proof is beyond the scope of this course:

   Cayley-Hamilton Theorem.

   Suppose A is an $(n \times n)$-square matrix.

   For each j, denote the coefficient of the j-th power term of $p_A(x)$ is c_j. (So $p_A(x) = c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1}+c_nx^n$ as polynomials.)

   Then $c_0I_n + c_1A + c_2A^2 + \cdots + c_{n-1}A^{n-1} + c_nA^n = \mathcal{O}_{n\times n}$.

10. **Proof of Theorem (4).**

    Suppose A is a diagaonalizable $(n \times n)$-square matrix.

    Then there are some non-singular $(n\times n)$-square matrix U and some real numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$ such that $U^{-1}AU = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

    For each $k = 1, 2, \cdots, n$, the number λ_k are eigenvalues of A. Then $p_A(\lambda_k) = 0$.

    Note that for each positive integer p, $U^{-1}A^pU = (U^{-1}AU)^p = (\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n))^p = \mathrm{diag}({\lambda_1}^p, {\lambda_2}^p, \cdots, {\lambda_n}^p)$.

    For each j, denote the coefficient of the j-th power term of $p_A(x)$ is c_j. (So $p_A(x) = c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1}+c_nx^n$ as polynomials.)

    We have

$$
\begin{aligned}
& U^{-1}(c_0I_n + c_1A + c_2A^2 + \cdots + c_{n-1}A^{n-1} + c_nA^n)U \\
= \;& c_0I_n + c_1U^{-1}AU + c_2U^{-1}A^2U + \cdots + c_{n-1}U^{-1}A^{n-1}U + c_nU^{-1}A^nU \\
= \;& c_0I_n + c_1\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) + c_2\,\mathrm{diag}({\lambda_1}^2, {\lambda_2}^2, \cdots, {\lambda_n}^2) \\
& + \cdots + c_{n-1}\,\mathrm{diag}({\lambda_1}^{n-1}, {\lambda_2}^{n-1}, \cdots, {\lambda_n}^{n-1}) + c_n\,\mathrm{diag}({\lambda_1}^n, {\lambda_2}^n, \cdots, {\lambda_n}^n) \\
= \;& \mathrm{diag}(p_A(\lambda_1), p_A(\lambda_2), \cdots, p_A(\lambda_n)) = \mathrm{diag}(0, 0, \cdots, 0) = \mathcal{O}_{n\times n}
\end{aligned}
$$

    Then $c_0I_n + c_1A + c_2A^2 + \cdots + c_{n-1}A^{n-1} + c_nA^n = \mathcal{O}_{n\times n}$.