Linear Algebra IV

Ji Yonghyeon, Bae Dongsung

February 19, 2026

We cover the following topics in this note.

- Eigenvectors and Diagonalization.
    * Hessian Matrix
    * Differential Equation
- TBA.

Notation.

- Let $\mathbb{F}$ is a field (typical cases: $\mathbb{R}$ or $\mathbb{C}$).
- Let V is a finite-dimensional $\mathbb{F}$-vector space.
- Let $T : V \to V$ is a linear operator.

Observation (Choosing a basis to simplify a linear map)**.** Let $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis of V such that $[T]_{\mathcal{B}}$ is a diagonal matrix:

$$[T]_{\mathcal{B}} = [T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}_{n\times n}, \quad \text{i.e.,} \quad T(\mathbf{v}_i) = d_i\mathbf{v}_i \text{ with } 1 \le i \le n.$$

Then T may be very complicated, but with respect to $[T]_{\mathcal{B}}$ it looks nice.

Eigenvector & Eigenvalue

Definition. Let $T : V \to V$ be $\mathbb{F}$-linear. A nonzero vector $\mathbf{v} \in V \setminus \{\mathbf{0}\}$ is an *eigenvector* of T if there exists $\lambda \in \mathbb{F}$ such that

$$T(\mathbf{v}) = \lambda\mathbf{v} \in V.$$

The scalar λ is called the **eigenvalue** corresponding to $\mathbf{v}$.

Remark 1. If $\mathbf{v} \neq 0$ and $T(\mathbf{v}) = \lambda\mathbf{v}$, then the one-dimensional subspace $\mathbb{F}\mathbf{v}$ satisfying

$$\begin{array}{rcccl} T|_{\mathbb{F}\mathbf{v}} & : & \mathbb{F}\mathbf{v} & \longrightarrow & \mathbb{F}\mathbf{v} \\ & & c\mathbf{v} & \longmapsto & \lambda(c\mathbf{v}) \end{array} \qquad (\because T(c\mathbf{v}) = cT(\mathbf{v}) = c\lambda\mathbf{v} = \lambda(c\mathbf{v})),$$

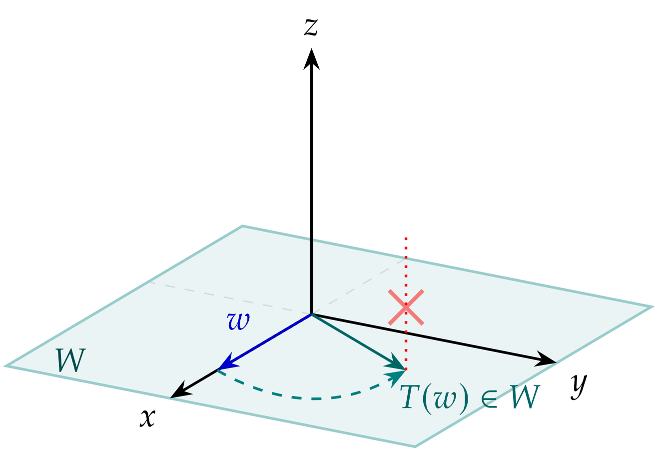
Equivalently, the restriction $T|_{\mathbb{F}\mathbf{v}}: \mathbb{F}\mathbf{v} \to \mathbb{F}\mathbf{v}$ acts as scalar multiplication by λ.

Remark 2. Let $T : V \to V$ be $\mathbb{F}$-linear. Let a subspace $W \leq V$ satisfy

$$T[W] \subseteq W \qquad (\iff \forall\, \mathbf{w} \in W,\ T(\mathbf{w}) \in W).$$

1. The restriction map
$$T|_W : W \to W, \qquad \mathbf{w} \mapsto T(\mathbf{w})$$
is a well-defined linear operator on W.

2. If $\dim V < \infty$ and $\mathcal{B} = (\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ is a basis of V such that $(\mathbf{w}_1, \ldots, \mathbf{w}_k)$ is a basis of W, then
$$[T]_{\mathcal{B}} = \begin{pmatrix} A & * \\ 0 & B \end{pmatrix},$$
where A represents $T|_W$ and B represents the induced map on V/W.



$$\begin{array}{c|c} & \begin{array}{ccc} e_1 \in W & e_2 \in W & \end{array} \\ \begin{array}{c} x \\ y \\ z \end{array} & \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{array}$$

Diagonalizability of Linear Operator

Definition. We say $T : V \to V$ is *diagonalizable* if $\exists$ a basis $\mathcal{B}$ of V such that $[T]_{\mathcal{B}}$ is diagonal.

Remark 3. T is diagonalizable if and only if V has a basis consisting of eigenvectors of T.
A diagonal matrix

$$\begin{bmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{bmatrix}$$

acts by scaling each basis vector $\mathbf{v}_i$ *independently*, and "scaling a nonzero vector" is the eigenvector condition $T(\mathbf{v}_i) = d_i \mathbf{v}_i$.

Example 1 (Hessian; Quadratic form diagonalization). Note that

- Single variable Taylor series:

$$f(x) = f(p) + \frac{1}{1!}f'(p)(x-p) + \frac{1}{2!}f''(p)(x-p)^2 + \cdots = \sum_{k=0}^{n} \frac{f^{(k)}(p)}{k!}(x-p)^k + R.$$

- Two variables Taylor series:

$$\begin{aligned} f(x,y) &= f(p,q) + \frac{1}{1!} f_x(p,q)(x-p) + \frac{1}{1!} f_y(p,q)(y-q) \\ &\quad + \frac{1}{2!} f_{xx}(p,q)(x-p)^2 + \frac{1}{1!} f_{xy}(p,q)(x-p)(y-q) + \frac{1}{2!} f_{yy}(p,q)(y-q)^2 + \cdots \end{aligned}$$

Let

$$X = \begin{bmatrix} x-p \\ y-q \end{bmatrix}, \qquad \nabla f = \begin{bmatrix} \frac{\partial}{\partial x} f \\ \frac{\partial}{\partial y} f \end{bmatrix}, \qquad H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}.$$

Then

$$\begin{aligned} f(x,y) &= \boxed{f(p,q)} + \boxed{\frac{1}{1!} \begin{bmatrix} f_x & f_y \end{bmatrix} \begin{bmatrix} x-p \\ y-q \end{bmatrix}} + \boxed{\frac{1}{2!} \begin{bmatrix} x-p & y-q \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \begin{bmatrix} x-p \\ y-q \end{bmatrix}} + R \\ &= \boxed{f(p,q)} + \boxed{\frac{1}{1!} \nabla f^T X} + \boxed{\frac{1}{2!} X^T H X} + R. \end{aligned}$$

Consider $f(x,y) = 2x^2 + 2xy + 2y^2$. Then $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$. Let

$$X = \begin{bmatrix} x \\ y \end{bmatrix} \quad \Rightarrow \quad f = X^T H X.$$

Eigenpairs:

$$\lambda_1 = 6 \Rightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \lambda_2 = 2 \Rightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad D = \mathrm{diag}(\lambda_1, \lambda_2) = \mathrm{diag}(6, 2), \qquad H = Q^T D Q.$$

Hence

$$f = X^T H X = X^T (Q^T D Q) X = (QX)^T D (QX).$$

Let $QX := V = \begin{bmatrix} u \\ v \end{bmatrix}$ then

$$u = \frac{x+y}{\sqrt{2}}, \qquad v = \frac{x-y}{\sqrt{2}}.$$

Then

$$f = V^T D V = \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = 6u^2 + 2v^2.$$

Here, intersection term (xy) disappeared.

Note. Let $A \in \mathrm{Mat}_n(\mathbb{F})$. The associated linear map $L_A : \mathbb{F}^n \to \mathbb{F}^n$ is given by $L_A(\mathbf{x}) = A\mathbf{x}$.

Diagonalizability of Matrix

Definition 1. A matrix $A \in \mathrm{Mat}_n(\mathbb{F})$ is *diagonalizable over* $\mathbb{F}$ if $\exists P \in \mathrm{GL}_n(\mathbb{F})$ and a diagonal matrix D such that

$$P^{-1}AP = D.$$

Eigenbasis and Similarity

Proposition 1. *Let $A \in \mathrm{Mat}_n(\mathbb{F})$.*

(i) If $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis of $\mathbb{F}^n$ consisting of eigenvectors of A, and $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, then $P^{-1}AP$ is diagonal with $P \in \mathrm{GL}_n(\mathbb{F})$.

(ii) Conversely, if $P^{-1}AP = D$ is diagonal, then the columns of P form an eigenbasis of A (with eigenvalues given by the diagonal entries of D).

Characteristic polynomial

Definition 2. For $A \in \mathrm{Mat}_n(\mathbb{F})$, the **characteristic polynomial** of A is

$$\chi_A(\lambda) := \det(A - \lambda I_n) \in \mathbb{F}[\lambda].$$

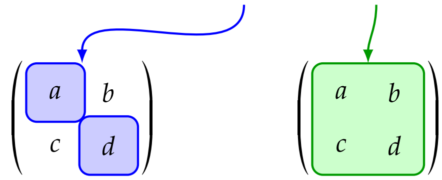
Eigenvalues are roots

Proposition 2. *A scalar $\lambda \in \mathbb{F}$ is an eigenvalue of A if and only if $\chi_A(\lambda) = 0$.*

Observation (Characteristic polynomial in 2×2). Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{Mat}_2(\mathbb{F})$. Then

$$\begin{aligned} \det(A - \lambda I_2) &= \det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \\ &= (a-\lambda)(d-\lambda) - bc \\ &= ad - (a+d)\lambda + \lambda^2 - bc \\ &= \lambda^2 - (a+d)\lambda + (ad - bc) \\ &= \lambda^2 - \mathrm{tr}(A)\lambda + \det(A). \end{aligned}$$

$$\chi_A(\lambda) = \lambda^2 - (\mathrm{tr}A)\,\lambda + \det(A)$$

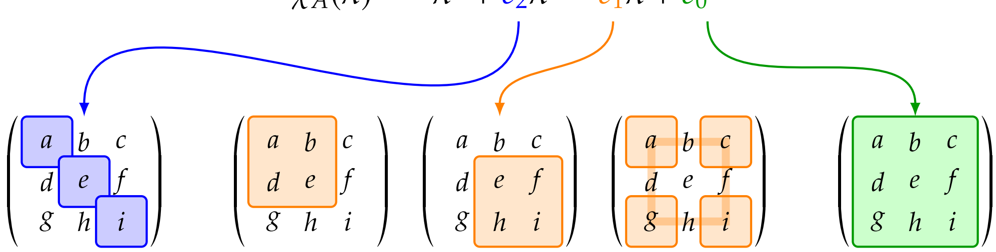


Observation (Characteristic polynomial in 3×3). Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \in \mathrm{Mat}_3(\mathbb{F})$. Then

$$\det(A - \lambda I_3) = \det \begin{bmatrix} a-\lambda & b & c \\ d & e-\lambda & f \\ g & h & i-\lambda \end{bmatrix} = (a-\lambda)\det\begin{bmatrix} e-\lambda & f \\ h & i-\lambda \end{bmatrix} - b\det\begin{bmatrix} d & f \\ g & i-\lambda \end{bmatrix} + c\det\begin{bmatrix} d & e-\lambda \\ g & h \end{bmatrix}$$

$$\begin{aligned} &= (a-\lambda)\big((e-\lambda)(i-\lambda) - fh\big) - b\big(d(i-\lambda) - fg\big) + c\big(dh - ge + g\lambda\big) \\ &= (a-\lambda)\big(ei - (e+i)\lambda + \lambda^2 - fh\big) - b\big(di - d\lambda - fg\big) + c\big(dh - ge + g\lambda\big) \\ &= \big(aei - a(e+i)\lambda + a\lambda^2 - afh\big) - \big(ei\lambda - (e+i)\lambda^2 + \lambda^3 - fh\lambda\big) \\ &\quad - \big(bdi - bd\lambda - bfg\big) + \big(cdh - cge + cg\lambda\big) \\ &= -\lambda^3 + (a+e+i)\lambda^2 - (ae + ai + ei - fh - bd - cg)\lambda + (aei - afh - bdi - bfg + cdh - cge) \\ &= -\lambda^3 + \mathrm{tr}(A)\lambda^2 - \left(\det\begin{bmatrix} a & b \\ d & e \end{bmatrix} + \det\begin{bmatrix} a & c \\ g & i \end{bmatrix} + \det\begin{bmatrix} e & f \\ h & i \end{bmatrix}\right)\lambda + \det(A). \end{aligned}$$

$$\chi_A(\lambda) = -\lambda^3 + c_2\lambda^2 - c_1\lambda + c_0$$



Note (Principal minors). A principal minor of a square matrix is the determinant of a principal submatrix, which is formed by selecting the same set of indices for both rows and columns.

Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$ and $\{i_1 < i_2 < \cdots < i_\ell\} \subset \{1, \ldots, n\}$. For each $\ell = 1, \ldots, n$, we define

$$E_\ell(A) := \sum_{1 \le i_1 < \cdots < i_k \le n} \det \begin{bmatrix} a_{i_1,i_1} & a_{i_1,i_2} & \cdots & a_{i_1,i_k} \\ a_{i_2,i_1} & a_{i_2,i_2} & \cdots & a_{i_2,i_k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_k,i_1} & a_{i_k,i_2} & \cdots & a_{i_k,i_k} \end{bmatrix},$$

so

$$\begin{aligned} E_1(A) &:= \sum_{i=1}^n \det \begin{bmatrix} a_{ii} \end{bmatrix} = \sum_{i=1}^n a_{ii} = \mathrm{tr}(A), \\ E_2(A) &:= \sum_{1 \le i < j \le n} \det \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}, \\ E_3(A) &:= \sum_{1 \le i < j < k \le n} \det \begin{bmatrix} a_{ii} & a_{ij} & a_{ik} \\ a_{ji} & a_{jj} & a_{jk} \\ a_{ki} & a_{kj} & a_{kk} \end{bmatrix}, \\ &\cdots, \\ E_n(A) &:= \det(A). \end{aligned}$$

Characteristic polynomial: Trace and Determinant as coefficients

Theorem 3. *Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$. Define the characteristic polynomial*

$$\chi_A(\lambda) := \det(A - \lambda I_n) = \sum_{i=0}^n c_i \lambda^i \in \mathbb{F}[\lambda].$$

Then $\chi_A(\lambda)$ is

$$\chi_A(\lambda) = (-1)^n \left(\lambda^n - E_1(A)\lambda^{n-1} + E_2(A)\lambda^{n-2} - \cdots + (-1)^n E_n(A) \right).$$

Proof. We use the mathematical induction on n.

(Base case $n = 1$) If $A = [a]$, then

$$\chi_A(\lambda) = \det[a - \lambda] = a - \lambda = (-1)^1 \lambda^1 + (-1)^0 a.$$

Thus $c_1 = -1 = (-1)^1$, $c_0 = a = \det(A)$, and $c_0 = (-1)^0\, \mathrm{tr}(A)$ is consistent since $\mathrm{tr}(A) = a$.

(Induction step) Assume the theorem holds for all $(n-1)\times(n-1)$ matrices over $\mathbb{F}$. Let $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{F})$ and set

$$M(\lambda) \coloneqq A - \lambda I_n.$$

Expand $\det(M(\lambda))$ by Laplace expansion along the first row:

$$\chi_A(\lambda) = \det(M(\lambda)) = \sum_{j=1}^{n} (-1)^{1+j} \,(a_{1j} - \lambda\delta_{1j})\, \det\big(M(\lambda)_{1j}\big), \tag{1}$$

where $M(\lambda)_{1j}$ denotes the $(n-1)\times(n-1)$ matrix obtained by deleting row 1 and column j.

Claim 1: $c_n = (-1)^n$ and $\deg \chi_A = n$. Observe that $\det(M(\lambda)_{1j})$ is a polynomial in λ of degree at most $n-1$. Moreover:

- If $j \neq 1$, then $(a_{11} - \lambda)$ does not appear, and the prefactor is a_{1j} (independent of λ). Hence the corresponding summand in (1) has degree $\leq n-1$.
- If $j = 1$, then the prefactor is $(a_{11} - \lambda)$, so the degree of that summand is $1 + \deg\det(M(\lambda)_{11})$.

Therefore the coefficient of λ^n can only come from the $j = 1$ summand:

$$(a_{11} - \lambda)\det(M(\lambda)_{11}).$$

Now identify $M(\lambda)_{11}$ explicitly. Deleting row 1 and column 1 removes the first diagonal position, so

$$M(\lambda)_{11} = A_{11} - \lambda I_{n-1},$$

where A_{11} is the $(n-1)\times(n-1)$ minor of A (delete row 1, column 1). By the induction hypothesis applied to A_{11},

$$\det(A_{11} - \lambda I_{n-1}) = (-1)^{n-1}\lambda^{n-1} + (\text{lower degree terms}).$$

Multiplying by $(a_{11} - \lambda)$, the λ^n term is

$$(a_{11} - \lambda)\Big((-1)^{n-1}\lambda^{n-1} + \cdots\Big) = (-\lambda)\cdot(-1)^{n-1}\lambda^{n-1} + \cdots = (-1)^n\lambda^n + \cdots.$$

Hence $c_n = (-1)^n$, and in particular $\deg \chi_A = n$.

Claim 2: $c_{n-1} = (-1)^{n-1}\operatorname{tr}(A)$. We extract the coefficient of λ^{n-1} from (1). Split the sum into the $j = 1$ term and the $j \neq 1$ terms.

(i) Contribution from $j = 1$. Write the expansion (by induction hypothesis) for the $(n-1)\times(n-1)$ matrix A_{11}:

$$\det(A_{11} - \lambda I_{n-1}) = (-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}\operatorname{tr}(A_{11})\,\lambda^{n-2} + \cdots.$$

Then

$$\begin{aligned}(a_{11}-\lambda)\det(A_{11}-\lambda I_{n-1}) &= a_{11}\Big((-1)^{n-1}\lambda^{n-1}+\cdots\Big)-\lambda\Big((-1)^{n-1}\lambda^{n-1}+(-1)^{n-2}\operatorname{tr}(A_{11})\lambda^{n-2}+\cdots\Big)\\ &= a_{11}(-1)^{n-1}\lambda^{n-1} \ + \ \underbrace{(-1)^{n-1}(-\lambda)\lambda^{n-1}}_{\text{degree } n} \ + \ \Big[-(-1)^{n-2}\operatorname{tr}(A_{11})\lambda^{n-1}\Big]+\cdots.\end{aligned}$$

Thus the coefficient of λ^{n-1} coming from $j=1$ equals

$$(-1)^{n-1}a_{11} \ - \ (-1)^{n-2}\operatorname{tr}(A_{11}) = (-1)^{n-1}\big(a_{11}+\operatorname{tr}(A_{11})\big).$$

(ii) Contribution from $j \neq 1$. For $j \neq 1$, the prefactor in (1) is a_{1j} (degree 0 in λ), so only the λ^{n-1} term of $\det(M(\lambda)_{1j})$ could contribute. But $\det(M(\lambda)_{1j})$ is a determinant of size $(n-1)\times(n-1)$ in which the diagonal contains only $(n-2)$ entries of the form $(\cdot - \lambda)$: indeed, deleting column $j \neq 1$ removes the diagonal position corresponding to index j while deleting row 1 removes index 1, so among indices $\{2,\ldots,n\}$ we are missing one diagonal index. Consequently,

$$\deg\det(M(\lambda)_{1j}) \le n-2 \qquad (j \neq 1),$$

and therefore the $j \neq 1$ summands contribute nothing to the λ^{n-1} coefficient.

Combining (i) and (ii),

$$c_{n-1} = (-1)^{n-1}\big(a_{11}+\operatorname{tr}(A_{11})\big).$$

Finally, since $\operatorname{tr}(A) = a_{11}+\operatorname{tr}(A_{11})$ (the trace is the sum of diagonal entries, and A_{11} contains exactly the remaining diagonal entries $a_{22},\ldots,a_{nn}$), we get

$$c_{n-1} = (-1)^{n-1}\operatorname{tr}(A).$$

Claim 3: $c_0 = \det(A)$**.** Evaluating at $\lambda = 0$ gives

$$c_0 = \chi_A(0) = \det(A - 0\cdot I_n) = \det(A).$$

(This step is compatible with induction, but does not require it.)

This completes the induction and proves all three coefficient identities.

Write A in block form by separating the first row and column:

$$A = \begin{pmatrix} a_{11} & r \\ c & B \end{pmatrix},$$

where $r \in \mathrm{Mat}_{1\times(n-1)}(\mathbb{F})$ is the first row with the first entry removed, $c \in \mathrm{Mat}_{(n-1)\times 1}(\mathbb{F})$ is the first column with the first entry removed, and $B \in \mathrm{Mat}_{n-1}(\mathbb{F})$ is the $(n-1)\times(n-1)$ lower-right block.

Then

$$A - \lambda I_n = \begin{pmatrix} a_{11} - \lambda & r \\ c & B - \lambda I_{n-1} \end{pmatrix}.$$

(1) Degree and leading coefficient. Expand $\det(A - \lambda I_n)$ along the first row:

$$\det(A - \lambda I_n) = (a_{11} - \lambda)\det(B - \lambda I_{n-1}) + \sum_{j=2}^{n} (-1)^{1+j} a_{1j} \det(M_{1j}(\lambda)),$$

where $M_{1j}(\lambda)$ is the $(n-1)\times(n-1)$ minor obtained by deleting row 1 and column j.

Key observation (degree-counting in matrix form):

- $\det(B - \lambda I_{n-1})$ has degree $n-1$ with leading term $(-1)^{n-1}\lambda^{n-1}$.
- For $j \geq 2$, the matrix $M_{1j}(\lambda)$ is obtained from $B - \lambda I_{n-1}$ by deleting one of its columns (corresponding to the deleted column j). Hence $M_{1j}(\lambda)$ contains at most $n-2$ diagonal entries of the form $(\cdot - \lambda)$, so *every* term in $\det(M_{1j}(\lambda))$ has degree at most $n-2$. Thus $\deg \det(M_{1j}(\lambda)) \leq n-2$.

Therefore the only source of a λ^n term is

$$(a_{11} - \lambda)\det(B - \lambda I_{n-1}),$$

and its λ^n coefficient is

$$(-\lambda) \cdot \left((-1)^{n-1}\lambda^{n-1}\right) = (-1)^n \lambda^n.$$

Hence $c_n = (-1)^n$ and $\deg \chi_A = n$.

(2) The λ^{n-1} coefficient and the trace. From the same expansion, the summation terms with $j \geq 2$ cannot contribute to λ^{n-1} because they have degree $\leq n-2$. Hence c_{n-1} comes solely from

$$(a_{11} - \lambda)\det(B - \lambda I_{n-1}).$$

Write the first two top-degree terms of $\det(B - \lambda I_{n-1})$ in the same convention:

$$\det(B - \lambda I_{n-1}) = (-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}\operatorname{tr}(B)\,\lambda^{n-2} + (\text{lower powers}).$$

(Here we are using the $(n-1)\times(n-1)$ case, i.e. induction on size.)

Multiply:

$$\begin{aligned}(a_{11}-\lambda)\det(B-\lambda I_{n-1}) &= a_{11}\Big((-1)^{n-1}\lambda^{n-1}+\cdots\Big)-\lambda\Big((-1)^{n-1}\lambda^{n-1}+(-1)^{n-2}\operatorname{tr}(B)\lambda^{n-2}+\cdots\Big)\\ &= \Big[(-1)^{n-1}a_{11}-(-1)^{n-2}\operatorname{tr}(B)\Big]\lambda^{n-1}+(\text{terms of degree}\neq n-1).\end{aligned}$$

Thus

$$c_{n-1}=(-1)^{n-1}a_{11}-(-1)^{n-2}\operatorname{tr}(B)=(-1)^{n-1}\big(a_{11}+\operatorname{tr}(B)\big).$$

But $\operatorname{tr}(A)=a_{11}+\operatorname{tr}(B)$ (trace is the sum of diagonal entries; B contains $a_{22},\dots,a_{nn}$), so

$$c_{n-1}=(-1)^{n-1}\operatorname{tr}(A).$$

(3) Constant term. Evaluating at $\lambda=0$ gives

$$c_0=\chi_A(0)=\det(A-0\cdot I_n)=\det(A).$$

□

References

[1] 수학의 즐거움, Enjoying Math. "수학 공부, 기초부터 대학원 수학까지, 31. 선형대수학 (h) 고유벡터와 행렬의 대각화 -1" YouTube Video, 29:46. Published November 06, 2019. URL: https://www.youtube.com/watch?v=RSOxa1rI_Kk.