Math 121B — Linear Algebra

Neil Donaldson

Fall 2022

Linear Algebra, Stephen Friedberg, Arnold Insel & Lawrence Spence, 4th Ed 2003, Prentice Hall.

Review from 121A

We begin by recalling a few basic notions and notations.

Vector Spaces Bold-face $\mathbf{v}$ denotes a *vector* in a *vector space* V over a *field* $\mathbb{F}$. A vector space is closed under *vector addition* and *scalar multiplication*

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V,\ \forall \lambda_1, \lambda_2 \in \mathbb{F}, \quad \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 \in V$$

Examples. Here are four (families of) vector spaces over the field $\mathbb{R}$.

- $\mathbb{R}^2 = \{x\mathbf{i} + y\mathbf{j} : x, y \in \mathbb{R}\} = \{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) : x, y \in \mathbb{R}\}$ is a vector space over the field $\mathbb{R}$.
- $P_n(\mathbb{R})$; *polynomials* with degree $\leq n$ and coefficients in $\mathbb{R}$
- $P(\mathbb{R})$; polynomials over $\mathbb{R}$ with any degree.
- $C(\mathbb{R})$; continuous functions from $\mathbb{R}$ to $\mathbb{R}$.

Linear Combinations and Spans Let $\beta \subseteq V$ be a subset of a vector space V over $\mathbb{F}$. A *linear combination* of vectors in β is any *finite* sum

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n$$

where $\lambda_j \in \mathbb{F}$ and $\mathbf{v}_j \in \beta$. The *span* of β comprises all linear combinations: this is a *subspace* of V.

Bases and Co-ordinates A set $\beta \subseteq V$ is a *basis* of V if it has two properties:

Linear Independence Any linear combination yielding the zero vector is trivial; for distinct $\mathbf{v}_j \in \beta$,

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_n\mathbf{v}_n = \mathbf{0} \implies \forall j, \lambda_j = 0$$

Spanning Set $V = \operatorname{Span}\beta$; *every* vector in V is a (*finite!*) linear combination of elements of β.

Theorem. *β is a basis of V $\iff$* every *$\mathbf{v} \in V$ is a* unique linear combination *of elements of β.*
The cardinality of all basis sets is identical; this is the dimension $\dim_{\mathbb{F}} V$.

Example. $P_2(\mathbb{R})$ has standard basis $\beta = \{1, x, x^2\}$: every degree ≤ 2 polynomial is a unique as a linear combination $p(x) = a + bx + cx^2$ and so $\dim P_2(\mathbb{R}) = 3$. The real numbers a, b, c are the *co-ordinates* of p with respect to β; the *co-ordinate vector* of p is written

$$[p]_\beta = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

Linearity and Linear Maps A function $\mathrm{T} : V \to W$ between vector spaces V, W over the *same field* $\mathbb{F}$ is *($\mathbb{F}$-)linear* if it respects the linearity properties of V, W

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V,\ \forall \lambda_1, \lambda_2 \in \mathbb{F}, \quad \mathrm{T}(\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2) = \lambda_1 \mathrm{T}(\mathbf{v}_1) + \lambda_2 \mathrm{T}(\mathbf{v}_2)$$

We write $\mathcal{L}(V, W)$ for the set (indeed vector space!) of linear maps from V to W: this is shortened to $\mathcal{L}(V)$ if $V = W$. An *isomorphism* is an invertible/bijective linear map.

Theorem. *If $\dim_{\mathbb{F}} V = n$ and β is a basis of V, then the co-ordinate map $\mathbf{v} \mapsto [\mathbf{v}]_\beta$ is an isomorphism of vector spaces $V \to \mathbb{F}^n$.*

Matrices and Linear Maps If V, W are finite-dimensional, then any linear map $\mathrm{T} : V \to W$ can be described using matrix multiplication.

Example. If $A = \begin{pmatrix} 2 & -1 \\ 0 & -1 \\ -4 & 3 \end{pmatrix}$, then the linear map $\mathrm{L}_A : \mathbb{R}^2 \to \mathbb{R}^3$ (*left-multiplication by A*) is

$$\mathrm{L}_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 0 & -1 \\ -4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ -y \\ 3y - 4x \end{pmatrix}$$

The linear map in fact *defines* the matrix A; we recover the columns of the matrix by feeding the standard basis vectors to the linear map.

$$\begin{pmatrix} 2 \\ 0 \\ -4 \end{pmatrix} = \mathrm{L}_A \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} -1 \\ -1 \\ 3 \end{pmatrix} = \mathrm{L}_A \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

More generally, if $\mathrm{T} \in \mathcal{L}(V, W)$ and $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\gamma = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ are bases of V, W respectively, then the *matrix of* T *with respect to β and γ* is

$$[\mathrm{T}]_\beta^\gamma = \Big([\mathrm{T}(\mathbf{v}_1)]_\gamma \ \cdots \ [\mathrm{T}(\mathbf{v}_n)]_\gamma \Big) \in M_{m \times n}(\mathbb{F})$$

whose j^{th} column is obtained by feeding the j^{th} basis vector of β to T and taking its co-ordinate vector with respect to γ. This fits naturally with the co-ordinate isomorphisms

$$\mathrm{T}(\mathbf{v}) = \mathbf{w} \iff [\mathrm{T}]_\beta^\gamma [\mathbf{v}]_\beta = [\mathbf{w}]_\gamma$$

There are two special cases when $V = W$:

- If $\beta = \gamma$, then we simply write $[\mathrm{T}]_\beta$ instead of $[\mathrm{T}]_\beta^\beta$.
- If $\mathrm{T} = \mathrm{I}$ is the identity map, then $Q_\beta^\gamma := [\mathrm{I}]_\beta^\gamma$ is the *change of co-ordinate matrix* from β to γ.

Being able to convert linear maps into matrix multiplication is a central skill in linear algebra. Test your comfort by working through the following; if everything feels familiar, you should consider yourself in a good place as far as pre-requisites are concerned!

Example. Let $\mathrm{T} : P_2(\mathbb{R}) \to P_1(\mathbb{R})$ be the linear map defined by *differentiation*

$$\mathrm{T}(a + bx + cx^2) = b + 2cx \tag{$*$}$$

The standard bases of $P_2(\mathbb{R})$ and $P_1(\mathbb{R})$ are, respectively, $\beta = \{1, x, x^2\}$ and $\gamma = \{1, x\}$. Observe that

$$[\mathrm{T}(1)]_\gamma = [0]_\gamma = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad [\mathrm{T}(x)]_\gamma = [1]_\gamma = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad [\mathrm{T}(x^2)]_\gamma = [2x]_\gamma = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$
$$\implies [\mathrm{T}]_\beta^\gamma = \Big([\mathrm{T}(1)]_\gamma \; [\mathrm{T}(x)]_\gamma \; [\mathrm{T}(x^2)]_\gamma \Big) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Written in co-ordinates, we see the original linear map $(*)$

$$\big[\mathrm{T}(a + bx + cx^2)\big]_\gamma = [\mathrm{T}]_\beta^\gamma [a + bx + cx^2]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} b \\ 2c \end{pmatrix} = [b + 2cx]_\gamma \tag{$\dagger$}$$

1. $\eta = \{1 + x, x + x^2, x^2 + 1\}$ is also a basis of $P_2(\mathbb{R})$. Show that

$$[\mathrm{T}]_\eta^\gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix}$$

2. As in $(\dagger)$ above, the matrix multiplication

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a + b \\ 2b + 2c \end{pmatrix}$$

   corresponds to an equation $[\mathrm{T}(p)]_\gamma = [\mathrm{T}]_\eta^\gamma [p]_\eta$ for some polynomial $p(x)$; what is $p(x)$ in terms of a, b, c?

3. Find the change of co-ordinate matrix Q_η^β and check that the matrices of T are related by

$$[\mathrm{T}]_\eta^\gamma = [\mathrm{T}]_\beta^\gamma Q_\eta^\beta$$

1 Diagonalizability & the Cayley–Hamilton Theorem

1.1 Eigenvalues, Eigenvectors & Diagonalization (Review)

Definition 1.1. Suppose V is a vector space over $\mathbb{F}$ and $\mathrm{T} \in \mathcal{L}(V)$. A *non-zero* $\mathbf{v} \in V$ is an *eigenvector* of T with *eigenvalue* $\lambda \in \mathbb{F}$ (together an *eigenpair*) if

$$\mathrm{T}(\mathbf{v}) = \lambda \mathbf{v}$$

For matrices, the eigenvalues/vectors of $A \in M_n(\mathbb{F})$ are precisely those of $\mathrm{L}_A \in \mathcal{L}(\mathbb{F}^n)$.

Suppose λ is an eigenvalue of T;

1. The *eigenspace* of λ is the nullspace $E_\lambda := \mathcal{N}(\mathrm{T} - \lambda \mathrm{I})$.
2. The *geometric multiplicity* of λ of the dimension $\dim E_\lambda$.

We say that T is *diagonalizable* if there exists a basis of eigenvectors; an *eigenbasis.*

We start by recalling a couple of basic facts, the first of which is easily proved by induction.

Lemma 1.2. *If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are eigenvectors corresponding to distinct eigenvalues, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is linearly independent.*

Moreover, if $\dim_{\mathbb{F}} V = n$ and $\mathrm{T} \in \mathcal{L}(V)$ has n distinct eigenvalues, then T is diagonalizable.

Eigenvalues and Eigenvectors in finite dimensions

If $\dim_{\mathbb{F}} V = n$ and ϵ is a basis, then the eigenvector definition is equivalent to a matrix equation

$$[\mathrm{T}]_\epsilon [\mathbf{v}]_\epsilon = \lambda [\mathbf{v}]_\epsilon$$

In such a situation, T being diagonalizable means $\exists \beta$ such that $[\mathrm{T}]_\beta$ is a *diagonal matrix*

$$[\mathrm{T}]_\beta = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Thankfully there is systematic way to find eigenvalues and eigenvectors in finite-dimensions:

1. Choose any basis ϵ of V and compute the matrix $A = [\mathrm{T}]_\epsilon \in M_n(\mathbb{F})$.
2. Observe that

$$\begin{aligned} \lambda \in \mathbb{F} \text{ is an eigenvalue} &\iff \exists [\mathbf{v}]_\epsilon \in \mathbb{F}^n \setminus \{\mathbf{0}\} \text{ such that } A[\mathbf{v}]_\epsilon = \lambda [\mathbf{v}]_\epsilon \\ &\iff \exists [\mathbf{v}]_\epsilon \in \mathbb{F}^n \setminus \{\mathbf{0}\} \text{ such that } (A - \lambda I)\,[\mathbf{v}]_\epsilon = \mathbf{0} \\ &\iff \det(A - \lambda I) = 0 \end{aligned}$$

   This last is a degree n polynomial equation whose roots are the eigenvalues.

3. For each eigenvalue λ_j, compute the eigenspace $E_{\lambda_j} = \mathcal{N}(\mathrm{T} - \lambda_j \mathrm{I})$ to find the eigenvectors. Remember that E_{λ_j} is a subspace of the original vector space V, so translate back if necessary!

Definition 1.3. The *characteristic polynomial of* $\mathrm{T} \in \mathcal{L}(V)$ is the degree-n polynomial

$$p(t) := \det(\mathrm{T} - t\mathrm{I})$$

The eigenvalues of T are precisely the solutions to the *characteristic equation* $p(t) = 0$.

Examples 1.4. 1. $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has characteristic polynomial $p(t) = t^2 + 1 = (t+i)(t-i)$. As a linear map $\mathrm{L}_A \in \mathcal{L}(\mathbb{R}^2)$, A has *no eigenvalues* and *no eigenvectors*!

As a linear map $\mathrm{L}_A \in \mathcal{L}(\mathbb{C}^2)$, we have two eigenvalues $\pm i$. Indeed

$$(A - i\mathrm{I})\mathbf{v} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \mathbf{v} \implies E_i = \operatorname{Span} \begin{pmatrix} i \\ 1 \end{pmatrix}$$

and similarly $E_{-i} = \operatorname{Span} \begin{pmatrix} -i \\ 1 \end{pmatrix}$. We therefore have an eigenbasis $\beta = \{ \begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} -i \\ 1 \end{pmatrix} \}$ (of $\mathbb{C}^2$), with respect to which

$$[\mathrm{L}_A]_\beta = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

2. Let $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$ be defined by

$$\mathrm{T}(f)(x) = f(x) + (x-1)f'(x)$$

With respect to the standard basis $\epsilon = \{1, x, x^2\}$, we have the non-diagonal matrix

$$A = [\mathrm{T}]_\epsilon = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix} \implies p(t) = \det(A - tI) = (1-t)(2-t)(3-t)$$

With three distinct eigenvalues, T is diagonalizable. To find the eigenvectors, compute the nullspaces:

$$\lambda_1 = 1: \quad \mathbf{0} = (A - \lambda_1 I)[\mathbf{v}_1]_\epsilon = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{pmatrix} [\mathbf{v}_1]_\epsilon \implies [\mathbf{v}_1]_\epsilon \in \operatorname{Span} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \implies E_1 = \operatorname{Span}\{1\}$$

$$\lambda_2 = 2: \quad A - \lambda_2 I = \begin{pmatrix} -1 & -1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{pmatrix} \implies [\mathbf{v}_2]_\epsilon \in \operatorname{Span} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \implies E_2 = \operatorname{Span}\{1 - x\}$$

$$\lambda_3 = 3: \quad A - \lambda_3 I = \begin{pmatrix} -2 & -1 & 0 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \implies [\mathbf{v}_3]_\epsilon \in \operatorname{Span} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \implies E_3 = \operatorname{Span}\{1 - 2x + x^2\}$$

Making a sensible choice of non-zero eigenvectors, we obtain an eigenbasis, with respect to which the linear map is necessarily diagonal

$$\beta = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \{1, 1-x, 1-2x+x^2\} = \{1, 1-x, (1-x)^2\}$$

$$\mathrm{T}\big(a + b(1-x) + c(1-x)^2\big) = a + 2b(1-x) + 3c(1-x)^2$$

$$[\mathrm{T}]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Conditions for diagonalizability of finite-dimensional operators

We now borrow a little terminology from the theory of polynomials.

Definition 1.5. Let $\mathbb{F}$ be a field and $p(t)$ a polynomial with coefficients in $\mathbb{F}$.

1. Let $\lambda \in \mathbb{F}$ be a root; $p(\lambda) = 0$. The *algebraic multiplicity* $\text{mult}(\lambda)$ is the largest power of $\lambda - t$ to divide $p(t)$. Otherwise said, there exists[1]some polynomial $q(t)$ such that

$$p(t) = (\lambda - t)^{\text{mult}(\lambda)} q(t) \quad \text{and} \quad q(\lambda) \neq 0$$

2. We say that $p(t)$ *splits* over $\mathbb{F}$ if it factorizes completely into linear factors; equivalently $\exists a, \lambda_1, \ldots, \lambda_k \in \mathbb{F}$ such that

$$p(t) = a(\lambda_1 - t)^{m_1} \cdots (\lambda_k - t)^{m_k}$$

When $p(t)$ splits, the algebraic multiplicities sum to the degree n of the polynomial

$$n = m_1 + \cdots + m_k$$

Of course, we are most interested when $p(t)$ is the *characteristic polynomial* of a linear map $\mathrm{T} \in \mathcal{L}(V)$. If such a polynomial splits, then $a = 1$ and $\lambda_1, \ldots, \lambda_k$ are necessarily the (distinct) eigenvalues of T.

Example 1.6. The field matters! For instance $p(t) = t^2 + 1 = (t - i)(t + i) = -(i - t)(-i - t)$ splits over $\mathbb{C}$ but not over $\mathbb{R}$. Its roots are plainly $\pm i$.

For the purposes of review, we state the main result; this will be proved in the next section.

Theorem 1.7. *Let V be finite-dimensional. A linear map $\mathrm{T} \in \mathcal{L}(V)$ is diagonalizable if and only if,*

1. *Its characteristic polynomial splits over $\mathbb{F}$, and,*
2. *The geometric and algebraic multiplicities of each eigenvalue are equal;* $\dim E_{\lambda_j} = \text{mult}(\lambda_j)$.

Example 1.8. The matrix $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ is easily seen to have eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 5$. Indeed

$$p(t) = (3 - t)^2(5 - t), \qquad \text{mult}(3) = 2, \qquad \text{mult}(5) = 1$$

$$E_3 = \text{Span}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad E_5 = \text{Span}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \dim E_3 = \dim E_5 = 1$$

This matrix is *non-diagonalizable* since $\dim E_3 = 1 \neq 2 = \text{mult}(3)$.

Everything prior to this *should* be review. If it feels very unfamiliar, revisit your notes from 121A, particularly sections 5.1 and 5.2 of the textbook.

[1]The existence follows from Descartes *factor theorem* and the *division algorithm* for polynomials.

Exercises 1.1 1. For each matrix over $\mathbb{R}$; find its characteristic polynomial, its eigenvalues/spaces, and its algebraic and geometric multiplicities; decide if it is diagonalizable.

(a) $A = \begin{pmatrix} 2&0&0&0 \\ 0&3&1&0 \\ 0&0&3&1 \\ 0&0&0&3 \end{pmatrix}$ (b) $B = \begin{pmatrix} -1&6&0&0 \\ -2&6&0&0 \\ 0&0&3&0 \\ 0&0&0&3 \end{pmatrix}$

2. Suppose A is a *real* matrix with eigenpair $(\lambda, \mathbf{v})$. If $\lambda \notin \mathbb{R}$ show that $(\overline{\lambda}, \overline{\mathbf{v}})$ is also an eigenpair.

3. Show that the characteristic polynomial of $A = \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$ does not split over $\mathbb{R}$. Diagonalize A over $\mathbb{C}$.

4. Give an example of a 2×2 matrix whose entries are *rational numbers* and whose characteristic polynomial splits over $\mathbb{R}$, but not over $\mathbb{Q}$.

5. Diagonalize $\mathrm{L}_C \in \mathcal{L}(\mathbb{C}^2)$ where $C = \begin{pmatrix} 2i & 1 \\ 2 & 0 \end{pmatrix}$.

6. If $p(t)$ splits, explain why

$$\det \mathrm{T} = \lambda_1^{\mathrm{mult}(\lambda_1)} \cdots \lambda_k^{\mathrm{mult}(\lambda_k)}$$

where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T.

7. Suppose $\mathrm{T} \in \mathcal{L}(V)$ is invertible with eigenvalue λ. Prove that λ^{-1} is an eigenvalue of T^{-1} with the same eigenspace E_λ. If T is diagonalizable, prove that T^{-1} is also diagonalizable.

8. If V is finite-dimensional and $\mathrm{T} \in \mathcal{L}(V)$, we may define $\det \mathrm{T}$ to equal $\det[\mathrm{T}]_\beta$, where β is any basis of V. Explain why the choice of basis does not matter; that is, if γ is any other basis of V, we have $\det[\mathrm{T}]_\gamma = \det[\mathrm{T}]_\beta$.

1.2 Invariant Subspaces and the Cayley–Hamilton Theorem

The proof of Theorem 1.7 is facilitated by a new concept, of which eigenspaces are a special case.

Definition 1.9. Suppose $\mathrm{T} \in \mathcal{L}(V)$. A subspace W of V is T-*invariant* if $\mathrm{T}(W) \subseteq W$. In such a case, the *restriction* of T to W is the linear map

$$\mathrm{T}_W : W \to W : \mathbf{w} \mapsto \mathrm{T}(\mathbf{w})$$

Examples 1.10. 1. The trivial subspace $\{\mathbf{0}\}$ and the entire vector space V are invariant for any linear map $\mathrm{T} \in \mathcal{L}(V)$.

2. Every eigenspace is invariant; if $\mathbf{v} \in E_\lambda$, then $\mathrm{T}(\mathbf{v}) = \lambda\mathbf{v} \in E_\lambda$.

3. Continuing Example 1.8, if $A = \left(\begin{smallmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{smallmatrix}\right)$ then $W = \operatorname{Span}\{\mathbf{i}, \mathbf{j}\}$ is an invariant subspace for the linear map L_A. Indeed

$$A(x\mathbf{i} + y\mathbf{j}) = (3x + y)\mathbf{i} + 3y\mathbf{j} \in W$$

W is an example of a *generalized eigenspace*; we'll study these properly at the end of term.

To prove our diagonalization criterion, we need to see how to factorize the characteristic polynomial. It turns out that factors of $p(t)$ correspond to T-invariant subspaces!

Example 1.11. $W = \operatorname{Span}\{\mathbf{i}, \mathbf{j}\}$ is an invariant subspace of $A = \left(\begin{smallmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{smallmatrix}\right) \in M_3(\mathbb{R})$. With respect to the standard basis, the restriction $[\mathrm{L}_A]_W$ has matrix $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right)$. The characteristic polynomial $p_W(t)$ of the restriction is plainly a factor of the whole,

$$p(t) = (1-t)(2-t)(3-t) = (2-t)p_W(t)$$

Theorem 1.12. *Suppose* $\mathrm{T} \in \mathcal{L}(V)$*, that* $\dim V$ *is finite and that* W *is a* T*-invariant subspace of* V*. Then the characteristic polynomial of the restriction* T_W *divides that of* T*.*

The proof simply abstracts the approach of the example.

Proof. Extend a basis β_W of W to a basis β of V. Since $\mathrm{T}(\mathbf{w}) \in \operatorname{Span}\beta_W$ for each $\mathbf{w} \in W$, we see that the matrix of $[\mathrm{T}]$ has block form

$$[\mathrm{T}]_\beta = \left(\begin{array}{c|c} A & B \\ \hline O & C \end{array}\right) \implies p(t) = \det(A - tI)\det(C - tI) = p_W(t)\det(C - tI)$$

where $p_W(t)$ is the characteristic polynomial, and $A = [\mathrm{T}_W]_{\beta_W}$ the matrix of the restriction T_W. ∎

Corollary 1.13. *If* λ *is an eigenvalue of* T*, then* $\mathrm{T}_{E_\lambda} = \lambda\mathrm{I}_{E_\lambda}$ *is a multiple of the identity, whence,*

1. *The characteristic polynomial of the restriction* T_{E_λ} *is* $p_\lambda(t) = (\lambda - t)^{\dim E_\lambda}$.

2. $p_\lambda(t)$ *divides the characteristic polynomial of* T*. In particular* $\dim E_\lambda \le \operatorname{mult}(\lambda)$.

We are now in a position to state and prove an extended version of Theorem 1.7.

Theorem 1.14. *Suppose* $\dim_{\mathbb{F}} V = n$ *and that* $\mathrm{T} \in \mathcal{L}(V)$ *has distinct eigenvalues* $\lambda_1, \ldots, \lambda_k$. *The following are equivalent:*

1. T *is diagonalizable.*
2. *The characteristic polynomial splits over* $\mathbb{F}$ *and* $\dim E_{\lambda_j} = \mathrm{mult}(\lambda_j)$ *for each* j*; indeed*
$$p(t) = p_{\lambda_1}(t) \cdots p_{\lambda_k}(t) = (\lambda_1 - t)^{\dim E_{\lambda_1}} \cdots (\lambda_k - t)^{\dim E_{\lambda_k}}$$
3. $\sum_{j=1}^{k} \dim E_{\lambda_j} = n$
4. $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$

Example 1.15. $A = \begin{pmatrix} 7 & 0 & -12 \\ 0 & 1 & 0 \\ 2 & 0 & -3 \end{pmatrix}$ is diagonalizable. Indeed $p(t) = (1-t)^2(3-t)$ splits, and we have

λ	1	3
$\mathrm{mult}(\lambda)$	2	1
E_λ	$\mathrm{Span}\left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$	$\mathrm{Span} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$
$\dim E_\lambda$	2	1

and $\mathbb{R}^4 = E_1 \oplus E_3$

With respect to the eigenbasis $\beta = \left\{ \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}$, the map is diagonal $[\mathrm{L}_A]_\beta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Proof. $(1 \Rightarrow 2)$ If T is diagonalizable with eigenbasis β, then $[\mathrm{T}]_\beta$ is diagonal. But then

$$p(t) = (\lambda_1 - t)^{m_1} \cdots (\lambda_k - t)^{m_k}$$

splits and $\sum \mathrm{mult}(\lambda_i) = n$. The cardinality n of an eigenbasis is *at most* $\sum \dim E_{\lambda_i}$ since every element is an (independent) eigenvector. By Corollary 1.13 ($\dim E_{\lambda_j} \le \mathrm{mult}(\lambda_j)$) we see that

$$n \le \sum \dim E_{\lambda_j} \le \sum \mathrm{mult}(\lambda_j) = n \implies \forall j, \dim E_{\lambda_j} = \mathrm{mult}(\lambda_j)$$

whence the inequalities are *equalities* with each pair equal $\dim E_{\lambda_j} = \mathrm{mult}(\lambda_j)$

$(2 \Rightarrow 3)$ $p(t)$ splits $\implies n = \sum \mathrm{mult}(\lambda_j) = \sum \dim E_{\lambda_j}$

$(3 \Rightarrow 4)$ Assume $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_j}$ exists.[2] If $(\lambda_{j+1}, \mathbf{v}_{j+1})$ is an eigenpair, then $\mathbf{v}_{j+1} \notin E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_j}$ for otherwise this would contradict Lemma 1.2.

By induction, $E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ exists; by assumption it has dimension $n = \dim V$ and therefore *equals* V.

$(4 \Rightarrow 1)$ For each j, choose a basis β_j of E_{λ_j}. Then $\beta := \beta_1 \cup \cdots \cup \beta_k$ is a basis of V consisting of eigenvectors of T; an eigenbasis. ∎

[2] Distinct eigenspaces have trivial intersection: $i_1 \neq i_2 \le j \implies E_{i_1} \cap E_{i_2} = \{\mathbf{0}\}$.

T-cyclic Subspaces and the Cayley–Hamilton Theorem

We finish this chapter by introducing a general family of invariant subspaces and using them to prove a startling result.

Definition 1.16. Let $\mathrm{T} \in \mathcal{L}(V)$ and let $\mathbf{v} \in V$. The T-*cyclic subspace* generated by $\mathbf{v}$ is the span

$$\langle \mathbf{v} \rangle = \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \mathrm{T}^2(\mathbf{v}), \ldots\}$$

Example 1.17. Recalling Example 1.10.3 Let $A = \left(\begin{smallmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{smallmatrix}\right)$, and $\mathbf{v} = \mathbf{i} + \mathbf{k}$. It is easy to see that

$$A\mathbf{v} = 3\mathbf{i} + 5\mathbf{k}, \quad A^2\mathbf{v} = 9\mathbf{i} + 25\mathbf{k}, \quad \ldots, \quad A^m\mathbf{v} = 3^m\mathbf{i} + 5^m\mathbf{k}$$

all of which lie in $\operatorname{Span}\{\mathbf{i}, \mathbf{k}\}$. Plainly this is the L_A-cyclic subspace $\langle \mathbf{i} + \mathbf{k} \rangle$.

The proof of the following basic result is left as an exercise.

Lemma 1.18. *$\langle \mathbf{v} \rangle$ is the smallest T-invariant subspace of V containing $\mathbf{v}$, specifically:*

1. *$\langle \mathbf{v} \rangle$ is T-invariant.*
2. *If $W \leq V$ is T-invariant and $\mathbf{v} \in W$, then $\langle \mathbf{v} \rangle \leq W$.*
3. *$\dim \langle \mathbf{v} \rangle = 1 \iff \mathbf{v}$ is an eigenvector of T.*

We were lucky in the example that the general form $A^m\mathbf{v}$ was so clear. It is helpful to develop a more precise test for identifying the dimension and a basis of a T-cyclic subspace.

Suppose a T-cyclic subspace $\langle \mathbf{v} \rangle = \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \mathrm{T}^2(\mathbf{v}), \ldots\}$ has finite dimension.[3] Let $k \geq 1$ be maximal such that the set

$$\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\}$$

is linearly independent.

- If k doesn't exist, the *infinite* linearly independent set $\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots\}$ contradicts $\dim \langle \mathbf{v} \rangle < \infty$.
- By the maximality of k, $\mathrm{T}^k(\mathbf{v}) \in \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\}$; by induction this extends to

$$j \geq k \implies \mathrm{T}^j(\mathbf{v}) \in \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\}$$

It follows that $\langle \mathbf{v} \rangle = \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\}$, and we've proved a useful criterion.

Theorem 1.19. *Suppose $\mathbf{v} \neq \mathbf{0}$, then*

$$\begin{aligned} \dim \langle \mathbf{v} \rangle = k &\iff \{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\} \text{ is a basis of } \langle \mathbf{v} \rangle \\ &\iff k \text{ is maximal such that } \{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\} \text{ is linearly independent} \\ &\iff k \text{ is minimal such that } \mathrm{T}^k(\mathbf{v}) \in \operatorname{Span}\{\mathbf{v}, \mathrm{T}(\mathbf{v}), \ldots, \mathrm{T}^{k-1}(\mathbf{v})\} \end{aligned}$$

[3] Necessarily the situation if $\dim V < \infty$, when we are thinking about characteristic polynomials.

Examples 1.20. 1. According to the Theorem, in Example 1.17 we need only have noticed

- $\mathbf{v} = \mathbf{i} + \mathbf{k}$ and $A\mathbf{v} = 3\mathbf{i} + 5\mathbf{k}$ are linearly independent.
- That $A^2(\mathbf{i} + \mathbf{k}) = 9\mathbf{i} + 25\mathbf{k} \in \text{Span}\{\mathbf{v}, A\mathbf{v}\}$.

We could then conclude that $\langle \mathbf{v} \rangle = \text{Span}\{\mathbf{v}, A\mathbf{v}\}$ has dimension 2.

2. Let $\mathrm{T}(p(x)) = 3p(x) - p''(x)$ viewed as a linear map $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$ and consider the T-cyclic subspace generated by the polynomial $p(x) = x^2$

$$\mathrm{T}(x^2) = 3x^2 - 2, \quad \mathrm{T}^2(x^2) = \mathrm{T}(3x^2 - 2) = 3(3x^2 - 2) - 6 = 9x^2 - 12, \quad \ldots$$

Observe that $\{x^2, \mathrm{T}(x^2)\}$ is linearly independent, but that

$$\mathrm{T}^2(x^2) = 9x^2 - 12 = -9x^2 + 6(3x^2 - 2) \in \text{Span}\{x^2, \mathrm{T}(x^2)\}$$

We conclude that $\dim \langle x^2 \rangle = 2$. An alternative basis for $\langle x^2 \rangle$ is plainly $\{1, x^2\}$.

We finish by considering the interaction of a T-cyclic subspace with the characteristic polynomial. Surprisingly, the coefficients of the characteristic polynomial and the linear combination coincide.

Continuing the Example, if $W = \langle x^2 \rangle$ and $\beta_W = \{x^2, \mathrm{T}(x^2)\} = \{x^2, 3x^2 - 2\}$, then

$$[\mathrm{T}_W]_{\beta_W} = \begin{pmatrix} 0 & -9 \\ 1 & 6 \end{pmatrix} \implies p_W(t) = t^2 - 6t + 9$$

Theorem 1.21. *Let $\mathrm{T} \in \mathcal{L}(V)$ and suppose $W = \langle \mathbf{w} \rangle$ has $\dim W = k$ with basis*

$$\beta_W = \{\mathbf{w}, \mathrm{T}(\mathbf{w}), \ldots, \mathrm{T}^{k-1}(\mathbf{w})\}$$

in accordance with Theorem 1.19, then

1. *If $\mathrm{T}^k(\mathbf{w}) + a_{k-1}\mathrm{T}^{k-1}(\mathbf{w}) + \cdots + a_0\mathbf{w} = \mathbf{0}$, then the characteristic polynomial of T_W is*

$$p_W(t) = (-1)^k \left(t^k + a_{k-1}t^{k-1} + \cdots + a_1 t + a_0 \right)$$

2. *$p_W(\mathrm{T}_W) = 0$ is the* zero map *on W.*

Proof. 1. This is an exercise.

2. Write $\mathrm{S} \in \mathcal{L}(V)$ for the linear map

$$\mathrm{S} := p_W(\mathrm{T}) = (-1)^k(\mathrm{T}^k + a_{k-1}\mathrm{T}^{k-1} + \cdots + a_0\mathrm{I})$$

Part 1 says $\mathrm{S}(\mathbf{w}) = \mathbf{0}$. Since S is a polynomial in T, it commutes with all powers of T:

$$\forall j, \ \mathrm{S}(\mathrm{T}^j(\mathbf{w})) = \mathrm{T}^j(\mathrm{S}(\mathbf{w})) = \mathbf{0}$$

Since S is zero on the basis β_W of W, we see that S_W is the zero function. ∎

With a little sneakiness, we can drop the W's in the second part of the Theorem and observe an intimate relation between a linear map and its characteristic polynomial.

Corollary 1.22 (Cayley–Hamilton). *If V is finite-dimensional, then $\mathrm{T} \in \mathcal{L}(V)$ satisfies its characteristic polynomial; $p(\mathrm{T}) = 0$.*

Proof. Let $\mathbf{w} \in V$ and consider the cyclic subspace $W = \langle \mathbf{w} \rangle$ generated by $\mathbf{w}$. By Theorem 1.12,

$$p(t) = q_W(t) p_W(t)$$

for some polynomial q_W. But the previous result says that $p_W(\mathrm{T})(\mathbf{w}) = \mathbf{0}$, whence

$$p(\mathrm{T})(\mathbf{w}) = \mathbf{0}$$

Since we may apply this reasoning to *any* $\mathbf{w} \in V$, we conclude that $p(\mathrm{T})$ is the zero function. ∎

Examples 1.23. 1. $A = \left(\begin{smallmatrix} 2 & 1 \\ 3 & 4 \end{smallmatrix}\right)$ has $p(t) = t^2 - 6t + 5$ and we confirm:

$$A^2 - 6A = \begin{pmatrix} 7 & 6 \\ 18 & 19 \end{pmatrix} - 6 \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} = -5I$$

It may seem like a strange thing to do for this matrix, but the characteristic equation can be used to calculate the inverse of A:

$$A^2 - 6A + 5I = 0 \implies A(A - 6I) = -5I \implies A^{-1} = \frac{1}{5}(6I - A) = \frac{1}{5} \begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix}$$

2. We use the Cayley–Hamilton Theorem to compute A^4 when

$$A = \begin{pmatrix} 2 & -1 & \frac{8}{3} \\ 0 & 1 & -6 \\ 0 & 0 & 2 \end{pmatrix}$$

The characteristic polynomial is

$$p(t) = (2 - t)^2 (1 - t) = 4 - 8t + 5t^2 - t^3$$

By Cayley–Hamilton,

$$\begin{aligned} A^4 &= AA^3 = A(5A^2 - 8A + 4I) \\ &= 5A^3 - 8A^2 + 4A = 5(5A^2 - 8A + 4I) - 8A^2 + 4A \\ &= 17A^2 - 36A + 20I = 17 \begin{pmatrix} 4 & -3 & \frac{50}{3} \\ 0 & 1 & -18 \\ 0 & 0 & 4 \end{pmatrix} - 36 \begin{pmatrix} 2 & -1 & \frac{8}{3} \\ 0 & 1 & -6 \\ 0 & 0 & 2 \end{pmatrix} + 20 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 16 & -15 & \frac{562}{3} \\ 0 & 1 & -90 \\ 0 & 0 & 16 \end{pmatrix} \end{aligned}$$

3. Recall Example 1.4.2, where the linear map $\mathrm{T}(f(x)) = f(x) + (x-1)f'(x)$ had

$$p(t) = (1-t)(2-t)(3-t) = -t^3 + 6t^2 - 11t + 6$$

By Cayley–Hamilton, $\mathrm{T}^3 = 6\mathrm{T}^2 - 11\mathrm{T} + 6\mathrm{I}$. You can check this explicitly, after first computing

$$\mathrm{T}^2(f(x)) = f(x) + 3(x-1)f'(x) + (x-1)^2 f''(x), \text{ etc.}$$

Cayley–Hamilton can also be used to simplify higher powers of T and even to compute the inverse!

$$\mathrm{I} = \frac{1}{6}(\mathrm{T}^3 - 6\mathrm{T}^2 + 11\mathrm{T}) \implies \mathrm{T}^{-1} = \frac{1}{6}(\mathrm{T}^2 - 6\mathrm{T} + 11\mathrm{I})$$
$$\implies \mathrm{T}^{-1}(f(x)) = f(x) - \frac{1}{2}(x-1)f'(x) + \frac{1}{6}(x-1)^2 f''(x)$$

Exercises 1.2

1. For the linear map $\mathrm{T} = \mathrm{L}_A : \mathbb{R}^3 \to \mathbb{R}^3$ where $A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$ find the T-cyclic subspace generated by the standard basis vector $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

2. Let $\mathrm{T} = \mathrm{L}_A$, where $A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ and let $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Compute $\mathrm{T}(\mathbf{v})$ and $\mathrm{T}^2(\mathbf{v})$. Hence describe the T-cyclic subspace $\langle \mathbf{v} \rangle$ and its dimension.

3. Given $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}$, find *two distinct* L_A-invariant subspaces $W \leq \mathbb{R}^4$ such that $\dim W = 3$.

4. Suppose that W and X are T-invariant subspaces of V. Prove that the sum

$$W + X = \{\mathbf{w} + \mathbf{x} : \mathbf{w} \in W, \mathbf{x} \in X\}$$

is also T-invariant.

5. Prove Lemma 1.18.

6. Give an example of an infinite-dimensional vector space V, a linear map $\mathrm{T} \in \mathcal{L}(V)$, and a vector $\mathbf{v}$ such that $\langle \mathbf{v} \rangle = V$.

7. Let $\beta = \{\sin x, \cos x, 2x \sin x, 3x \cos x\}$ and $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x} \in \mathcal{L}(\operatorname{Span}\beta)$. Plainly the subspace $W := \operatorname{Span}\{\sin x, \cos x\}$ is T-invariant. Compute the matrices $[\mathrm{T}]_\beta$ and $[\mathrm{T}_W]_{\beta_W}$ and observe that

$$p(t) = \big(p_W(t)\big)^2$$

8. Verify explicitly that $A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$ satisfies its characteristic polynomial.

9. Check the details of Example 1.23.3 and evaluate T^4 as a linear combination of I, T and T^2. In particular, check the evaluation of $\mathrm{T}^{-1}(f(x))$.

10. Suppose a, b are constants with $a \neq 0$ and define $\mathrm{T}(f(x)) = af(x) + bf'(x)$.

    (a) Find an expression for the inverse $\mathrm{T}^{-1}(f(x))$ if $\mathrm{T} \in \mathcal{L}(P_1(\mathbb{R}))$

    (b) Find an expression for the inverse $\mathrm{T}^{-1}(f(x))$ if $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$

    Your answers should be written in terms of f and its derivatives.

11. Let $\mathrm{T}(f)(x) = f'(x) + \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t$ be a linear map $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$.

    (a) Find the characteristic polynomial of T and identify its eigenspaces. Is T diagonalizable?

    (b) Find $a, b, c \in \mathbb{R}$ such that $\mathrm{T}^3 = a\mathrm{T}^2 + b\mathrm{T} + c\mathrm{I}$.

    (c) What are $\dim \mathcal{L}(P_2(\mathbb{R}))$ and $\dim \operatorname{Span}\{\mathrm{T}^k : k \in \mathbb{N}_0\}$? Explain.

12. If $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ has non-zero determinant, use the Cayley–Hamilton Theorem to obtain the usual expression for A^{-1}.

13. Recall Examples 1.10.3, 1.17, and 1.20.1 with $A = \left(\begin{smallmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{smallmatrix}\right)$.

    (a) If $\mathbf{v} = \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$, show that $\det(\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}) = -4y^2z$

    (b) Hence determine *all* L_A-cyclic subspaces of $\mathbb{R}^3$.

14. (a) Consider Example 1.20.2 where $\mathrm{T} \in \mathcal{L}(P_2(\mathbb{R}))$ is defined by $\mathrm{T}(p(x)) = 3p(x) - p''(x)$. Prove that all T-cyclic subspaces have dimension ≤ 2.

    (b) What if we instead consider $\mathrm{S} \in \mathcal{L}(P_2(\mathbb{R})$ defined by $\mathrm{S}(p(x)) = 3p(x) - p'(x)$?

15. We prove part 1 of Theorem 1.21.

    (a) Explain why the matrix of T_W with respect to the basis β_W is

$$[\mathrm{T}_W]_{\beta_W} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & & 0 & -a_1 \\ 0 & 1 & 0 & & 0 & -a_2 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & & 0 & -a_{k-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{k-1} \end{pmatrix} \in M_k(\mathbb{F})$$

    (b) Compute the characteristic polynomial $p_W(t) = \det\big([\mathrm{T}_W]_{\beta_W} - tI_k\big)$ by expanding the determinant along the first row.